Is Affirmative Action in Employment Still Effective in the 21st Century?

Noriko Amano-Patiño[*] Zara Contractor[†] Julian Aramburu[‡]

Please click here for the latest version

Abstract

We reassess the most widespread employment-based affirmative action regulation in the US, Executive Order 11246, which increased representation of Black workers in the late 1900s. We find no positive impact of the regulation on the share of Black workers in firms regulated from 2001 to 2014. These results extend to other protected groups, and to earnings and worker flows. While firms entering regulation are positively selected, we show the regulation still has scope for continued impacts. Analyzing audited establishments, we find evidence of a disconnect between complying with the law as written versus implementing meaningful changes that impact workforce outcomes.

JEL-codes: J15, J23, J31, J71, J78, K31.
Keywords: Executive Order 11246, Affirmative Action in Employment, Racial Disparities.

Acknowledgments: **(1)** We thank Joseph Altonji, Jaime Arellano-Bover, Oscar Branson, Rosa Ferrer, Lucas Finamor, Daniel Giraldo, Sebastian Heise, Rebecca McKibbin, Costas Meghir, Conrad Miller, Javier Miranda, Caitlin Myers, Cormac O'Dea, Vitor Possebom, and Yanos Zylberberg for their insights and helpful discussions, which significantly enhanced our project. We also thank conference attendees at BAEM 2024, BSF 2024, and EEA-ESEM 2023 for their valuable comments and suggestions. Yidi Wu and Adam Hunt provided excellent research assistance. **(2)** We are grateful to Jiashen You, Justin West, and David Bowden at the EEOC for supporting this project and to Shirley Wilcher, Jeffery Lewis, Julia Achee, Zoe Ann Whitley, and Brad Wiltshire for their time, insights, and expertise. **(3)** We thank the Keynes Fund for generously funding this project.
Disclaimers: **(1)** This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by National Science Foundation Grants SES-9978093, SES-0339191, and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. Work on this article took place at the Yale Federal Statistical Research Data Center. **(2)** Julian Aramburu's work was conducted prior to his tenure at Amazon in April 2022. This paper is not sponsored, endorsed, or associated with Amazon or its subsidiaries or affiliates. **(3)** Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau, the EEOC, or Amazon. All results have been reviewed to ensure no confidential information is disclosed.

[*]University of Cambridge, noriko.amanopatino@econ.cam.ac.uk.
[†]Middlebury College, zcontractor@middlebury.edu.
[‡]Amazon, aramburu.julian@gmail.com

1 Introduction

The most widespread US affirmative action policy in employment, Executive Order 11246, played a crucial role in combatting racial disparities in the US labor market in the late 1900s. The policy requires firms with sizeable federal contracts to identify and address any underrepresentation of racial minorities and women in their workforce. The literature finds that regulated firms — which employ approximately a fifth of the workforce (Farrell, 2020) — increased their share of Black workers upon becoming regulated (Miller, 2017; Kurtulus, 2016, 2012; Rodgers and Spriggs, 1996; Leonard, 1984; Smith and Welch, 1984; Heckman and Wolpin, 1976; Goldstein and Smith, 1976; Ashenfelter and Heckman, 1974; Burman, 1973), with the latest of these studies finding increases that were enough to significantly reduce or even close the Black-white jobless gap in regulated firms (Miller, 2017).[1] Notably, the regulation achieved these gains without imposing quotas, and the gains persisted even after firms were no longer regulated (Miller, 2017; Kurtulus, 2016, 2012). Fifty-nine years after being established, this regulation continues to be a subject of policy debate[2] and features prominently in public discourse.[3]

So far, the literature on Executive Order 11246 has focused on its impact in the late 20th century, with the most recent studies examining data up to 2004. In this paper, we revisit this question for the years 2001–2014 and find a significant change. We find no evidence of newly regulated firms increasing representation of Black workers. Our estimates are generally insignificant and, in fact, trend negative. Considering additional workforce outcomes (i.e., average earnings, hires, and worker retention) and extending our analysis to other racial minorities and women, who are also covered by the regulation,[4] we find similarly small effects.

Our findings could be explained by firms entering the regulation status already hiring minority workers at representative levels or by a lack of compliance with the regulation. Although federal contractors, including new contract recipients, tend to hire more minorities than other

[1] Miller (2017) finds that the share of Black workers increases by 0.88 p.p. 5 years post-regulation and "a 0.8 to 1.3 percentage point increase in the black share of the US workforce would eliminate the black-white jobless gap over this period" in regulated firms p. 2.

[2] For example, Executive Order 13950, signed in 2020, prohibited certain types of diversity and inclusion training, directly affecting strategies allowed under the regulation we study, Executive Order 11246. This order was revoked by President Biden in 2021.

[3] Most recently because of the potential consequences of the Supreme Court decision on college admissions for federal contractors' diversity initiatives. See for example https://www.nytimes.com/2023/06/30/business/economy/hiring-affirmative-action.html.

[4] Executive Order 11246 was amended to include as protected characteristics "sex" in 1967, "sexual orientation" in 1998, and " gender identity" in 2014. Race and color were included since its inception in 1965. Although the term Hispanic was created in 1977 by the Office of Management and Budget, people with Spanish surnames were considered a minority covered by the regulation from the onset.

firms,[5] there is significant heterogeneity in employment disparities among them: approximately half of the new federal contractors continue to underrepresent minority workers. This indicates that Executive Order 11246 still has scope to increase minority representation. However, we find no positive effect of the policy on these firms, pointing to issues with compliance. Through an analysis of audited establishments, we find suggestive evidence that the key issue is a disconnect between complying with the law as written versus implementing meaningful changes that impact workforce outcomes.

We now go into the details of our analysis, starting with key aspects of the regulation we study. Executive Order 11246 applies to firms holding a contract or subcontract with the federal government (henceforth *federal contractors*). Federal contractors employing at least 50 workers and holding contracts exceeding $50,000 annually are subject to *affirmative action*, which means they must make a good-faith effort to hire and promote minorities and women at rates representative of the local qualified labor force. They must create an affirmative action plan (henceforth *AAP*) for each of their establishments, which involves identifying job groups in which protected workers are underrepresented relative to the local qualified labor force and implementing strategies to correct this imbalance. These strategies cannot involve quotas; instead, they focus on recruiting qualified candidates and improving the screening processes in place.[6] We refer to employment-based affirmative action as mandated by Executive Order 11246 as *affirmative action* throughout the paper.

To study the impact of affirmative action on establishment workforce outcomes, we link two key data sources. We combine workers' job histories from the Longitudinal Employer-Household Dynamics (LEHD) database[7] with their employers' federal contractor status from the Federal Procurement Data System (FPDS), creating a novel linked employer-employee panel dataset containing information on affirmative action regulation status. Our resulting

[5]This has been the case since the 1980s: Rodgers and Spriggs (1996) find that, except for Hispanic women, there were more minority workers among federal contractors than among non-contractor establishments (1982–1992). The average shares of minorities reported by Kurtulus (2016) suggest this continues to be the case (1973–2003).

[6]Separate regulations govern affirmative action for veterans (VEVRAA) and workers with disabilities (Section 503), which are beyond the scope of this paper. VEVRAA stands for the Vietnam Era Veterans' Readjustment Assistance Act of 1974, as amended, which requires contractors or subcontractors with a federal contract valued at $100,000 or more to take affirmative action to employ and advance in employment qualified covered veterans. Section 503 of the Rehabilitation Act of 1973 prohibits contractors or subcontractors with a federal contract valued at $10,000 or more from discriminating in employment against individuals with disabilities and requires employers to take affirmative action to recruit, hire, promote, and retain these individuals.

[7]We have access to the LEHD for 18 states and Washington, D.C. The states in our study are Arizona, Arkansas, Colorado, Delaware, Hawaii, Indiana, Iowa, Kansas, Maine, Maryland, Montana, Nevada, New Mexico, Oklahoma, Pennsylvania, Tennessee, Washington and Wisconsin. In Appendix E we show that our results are not driven by the set of states.

dataset has several advantages over data used by previous studies.[8] Notably, we do not rely on firms' self-reported regulation status, which has been found to be under-reported (see Appendix Table E.2).[9] Furthermore, the LEHD data let us study outcomes such as earnings and worker flows on a large scale for the first time.

We estimate the effects of becoming regulated by affirmative action with an event study design that compares workforce outcomes five years before and after firms receive a sizeable federal contract and thus become subject to this regulation. The workforce outcomes we study are the share of Black employees in the total workforce as well as in new hires, their retention rate, and their quarterly earnings and equivalent outcomes for other minorities and for women. Our analysis leverages time variation within establishments, supplemented by comparing federal contractors with contracts above $50,000 to those with contracts below this threshold.[10] In our main specification, we control for establishment size and age, weight observations by employment, and cluster standard errors at the firm level. We use the estimator proposed by Callaway and Sant'Anna (2021) to avoid biases from the staggered treatment across firms. We can identify the causal effect of affirmative action under the identifying assumption that becoming a federal contractor impacts the outcomes we study through, and only through, the onset of affirmative action. We will discuss how we deal with potential threats to identification after stating our results.

We first examine how the share of Black workers in an establishment evolves around the onset of the regulation. In contrast to previous findings, we find no evidence of a positive effect of affirmative action on this outcome. Our estimates are quantitatively small, and trend negative: the share of Black workers decreases by 0.03 percentage points (p.p.) in the year of regulation and by 0.25 p.p. five years post-regulation. The negative trend may be due to minority workers being hesitant to work in firms with affirmative action due to concerns of tokenism, as found by

[8]The most commonly used data source (EEO-1 reports) is comprised by establishment self-reported information, including workforce composition and 'affirmative action' status. The reports do not ask directly about affirmative action. Instead, they ask whether the company or any of its establishments have 50 or more employees and have a federal contract (or subcontract) of $50,000 or more, which implies they are regulated by affirmative action.

[9]Numerous posts from employment law consulting firms suggest possible discrepancies between an establishment's actual and perceived federal contractor status. See, for example, https://ogletree.com/insights-resources/blog-posts/i-think-my-company-is-a-federal-contractor-and-has-regulatory-obligations-but-where-can-i-look-to-search-for-that-information/, or https://www.littler.com/files/wp_covered_fed_gov_contractors_8-3-15_julie_2.pdf, or https://hortonpllc.com/what-is-executive-order-11246/.

[10]We exclude firms with fewer than 50 employees to prevent confounding the effects of affirmative action with those of the many other policies that apply to firms with 50 or more employees, such as the Family and Medical Leave Act.

the large-scale natural field experiment in Leibbrandt and List (2018).[11]

In contrast to our findings, for the years 1978–2004, Miller (2017) finds the regulation increased the share of Black workers by a statistically significant 0.1 p.p. in the first year of regulation, rising to 0.88 p.p. in five years. We show that the discrepancy between our results and the previous literature is explained almost fully by the time period studied rather than by any differences in the data or analysis. To do so, we replicate Miller's (2017) study using his data source and codes, changing *only* the time period to match ours. We obtain insignificant estimates comparable to our main results. In fact, we show that the regulation's effectiveness remained similarly strong until as recently as 1990–2005 and diminished steadily from then on.

Since an establishment may take time to adjust its workforce composition, the regulation could lead to hiring more or firing fewer Black workers without showing positive changes to the overall stock of workers. However, we find no positive impact on the share of Black workers among new hires or on the retention of Black workers. We test the regulation's impact on retention rates and find that Black workers are *less* likely to remain in treated firms; this result is small in magnitude but statistically significant. Affirmative action stipulates efforts should be made to increase diversity through promotions; we do not directly observe promotions, but we can estimate the regulation's impact on promotions by measuring Black workers' average yearly earnings, which would rise if they were promoted more frequently. Again, we find insignificant effects. Furthermore, our results are not particular to Black workers; we also find no positive impact on outcomes for Hispanics, all racial minorities as a group, or women.

Next, we consider threats to identification that could cause us to miss any true positive effect. We argue that our results are not driven by countervailing effects from other aspects of becoming a federal contractor, like increased labor demand. Although increased labor demand could theoretically lower the minority worker share, potentially offsetting a positive affirmative action effect, Aizer et al. (2020) show — in the context of World War II — that an increase in labor demand due to government contracts tends to *increase* minority hiring. Additionally, we supplement our event study analysis with a regression discontinuity design that separates the impact of the regulation from other aspects of receiving a federal contract. We find no significant impact of the regulation when comparing firms with similarly sized contracts right above and below the $50,000 threshold for affirmative action.

We then turn to consider that firms could enter our sample indirectly regulated, either through sub-contracts or other state-level affirmative action policies. This would indeed attenu-

[11] Leibbrandt and List (2018) do not directly study affirmative action. The authors study the impact of Equal Employment Opportunity statements in job ads on the application rates of minorities. They find that these statements *deter* minorities from applying.

ate our estimates of the regulations' effects but cannot explain why our estimates are *negative*. Moreover, even if firms entered the sample indirectly regulated, they would become subject to enforcement audits only upon entering our sample — subcontractors and employers subject to other policies are not subject to audits. In addition, we would still expect to see impacts on establishments that enter our sample with minority employment shares below reference rates, which we do not find.[12] Our supplementary regression discontinuity design also mitigates this concern by finding similar results without relying on variation from changes in regulation status.

We next consider whether positive effects on our control group could mask the regulation's impact. Our control group includes firms with federal contracts below the $50,000 threshold, which are not regulated by affirmative action. Notably, federal contractors with a contract value of $10,000 or more are subject to minimal provisions such as displaying equal opportunity posters and including equal opportunity taglines in job adverts. We show that the positive impacts of these minimal provisions in the control group do not drive our results by running an event study around the timing of becoming a federal contractor with a contract value above $10,000. Lastly, we consider selection into contracting. Once again, focusing on the discontinuity at $50,000 in contract value, we show that there is no bunching in contracts below $50,000 and that pre-contract outcomes are continuous across the threshold, suggesting that firms are not strategically choosing smaller contracts to avoid the regulation.

Having established that affirmative action regulation no longer shows positive effects in our time period, we turn to understanding why. Does the policy no longer have scope for effects, or is the lack of effects driven by a lack of compliance? We argue that our results are not driven by most federal contractors hiring minorities at rates representative of local populations. We estimate the share of local, qualified minority workers for each establishment using the ACS. While regulated firms have more diverse workforces on average, many still hire minorities at rates substantially below the local qualified averages. We show that our results hold even for establishments that underrepresent minorities when receiving a federal contract, which would be expected to identify and address these disparities in their AAP. Since we do not see evidence of shrinking disparities in these firms post-regulation, this suggests issues with compliance with the regulation.

To further investigate compliance issues; we turn to data on enforcement audits. The Office of Federal Contract Compliance Programs (OFCCP) enforces Executive Order 11246 by auditing approximately 1% of regulated establishments each year. If the absence of positive effects

[12]The *reference rates* we refer to here are the EEO-Tabulations which are a reference point for federal contractors to assess representation. These are not quotas.

were driven by firms perceiving a low risk of being audited, we would expect to see positive effects following audits, particularly for establishments found to be non-compliant. To test this, we augment our data with information on audited firms and find that audits have no positive impact on any of our workforce outcomes, even if the firm was found to be non-compliant. This suggests that simply increasing the number of audits would not be enough to restore the regulation's positive and significant effects of the late 1900s.

On average, 80% of audited establishments are found to be in compliance. For the remaining 20%, the most common outcome is a conciliatory agreement that may include provisions such as back pay.[13] However, passing an audit does not directly imply the regulation has effectively reduced any disparities in the firm. A report by the Government Accountability Office in 2016 (GAO, 2016) reveals that 85% of audited establishments require extensions to submit their AAPs after receiving an audit notice, suggesting they may not have updated AAPs readily available until they are notified of an impending audit. When firms do comply with the AAP requirement, they have considerable latitude in how they measure minority representation internally and in the local labor force, leaving room for inaccuracies in identifying areas of underrepresentation. Even where they do identify areas of underrepresentation and set strategies to address them, these strategies will not necessarily be effective, especially if they do not address the cause of the disparity. For example, a strategy to attract minority applicants may not improve outcomes if the disparity is caused by a biased screening algorithm, a prejudiced hiring manager, or if applicants are hesitant to apply due to a negative reputation of the firm.

Together, our findings underscore the challenges of effectively implementing and enforcing affirmative action policies in the current labor market context.

The rest of the paper is organized as follows. In Section 2, we review the related literature. We present the details and institutional background of Executive Order 11246 in Section 3 and the details of our data in Section 4. In Sections 5 and 6, we present our empirical strategies and results, which we discuss and conclude in Section 7.

[13] By *back pay* we mean compensation for wages and benefits that were unlawfully withheld or denied to workers due to discriminatory practices. It may cover work already performed or work that could have been performed had the worker not been subject to discrimination. These types of financial penalties are rare, occurring in less than 2% of cases of noncompliance.

2 Prior evidence on Executive Order 11246

We contribute to a long-standing body of literature evaluating the effects of Executive Order 11246 on employment. The first empirical evidence dates to the 1970s, shortly after the introduction of Executive Order 11246 in 1965.[14] Burman (1973), Ashenfelter and Heckman (1974), Goldstein and Smith (1976), Heckman and Wolpin (1976), Smith and Welch (1984) and Leonard (1984) find that affirmative action was effective in increasing the employment share of Black men in the 1970s, with modest effects for men of other protected groups. Goldstein and Smith (1976), Heckman and Wolpin (1976), Smith and Welch (1984) and Leonard (1984) find mixed though generally positive effects for of the policy on women.

The Reagan Administration (1981–1989) brought numerous changes to the regulation, including attempts to rescind the Executive Order.[15] During this time, the evidence on the regulation's efficacy is mixed: Rodgers and Spriggs and Miller (2017) find continued effects while Kurtulus (2016), and Leonard (1990) suggest little to no impact.

Kurtulus (2012), Kurtulus (2016), and Miller (2017) study the long-term effects of affirmative action over similar periods spanning from the early 1970s up 2004. These are the papers closest to ours as they use event study specifications to account for the dynamic impacts of the regulation; that is, firms that used to be regulated may retain positive impacts. Kurtulus (2012, 2016) finds positive impacts of the policy across protected groups, particularly in high wage occupations. Miller (2017) finds the regulation increased the share of Black employees at regulated firms by 0.88 p.p. five years after first regulation — which would almost eliminate the black-white jobless gap in regulated firms over this period (Miller, 2017) — and that these increased shares persisted even after a firm was no longer a federal contractor.

Our paper is the first to study the impact of Executive Order 11246 during a more recent time period, 2001–2014, show that the previous positive results no longer hold and explore why. Another key contribution of our paper is to create a dataset linking information on federal contracts with the administrative matched employer-employee data. The detailed data on federal contracts allows us to infer regulation status without relying on self-reported measures, which

[14] Holzer and Neumark (2000) comprehensively review the literature produced until 2000. Farrell (2020) provides an excellent overview of Executive Order 11246 and its enforcement by each presidential administration.

[15] Anderson (1996) and Leonard (1996) documented a sharp decline in federal enforcement activity in the 1980s during the Reagan Administration (1981–1989). Although the efforts to rescind the Executive Order failed, new regulations raised the threshold for coverage to 250,000 per contract, reducing the number of contractors required to adopt an AAP. Additionally, Leonard (1996) documents a sharp decrease in administrative complaints filed, back-pay awards, and the disappearance of debarment (which were already rare) after 1980.

have been found to be under-reported by firms.[16] The administrative data allow us to study the impacts of Executive Order 11246 on additional outcomes beyond total employment: composition of new hires, retention rates and earnings.[17] We build on papers studying audits of the regulation, Goldstein and Smith (1976) and Leonard (1984), by incorporating audit outcomes and accounting for dynamic effects in an event study framework. Furthermore, we are the first paper to study how establishment-level representation compares to the local qualified levels, as specified by the regulation.

Kline et al. (2022) and Kline et al. (2024) experimentally analyze discrimination in hiring, revealing racial gaps in callback rates. While these papers do not directly study affirmative action, they show heterogeneity by firms' federal contractor status. They find federal contractors have, on average, smaller racial callback gaps. However, several federal contractors were among the most discriminatory firms in their sample. This supports our finding that firms regulated by affirmative action still have scope to combat racial disparities.

3 Institutional Background

Executive Order 11246 (EO 11246 henceforth), which was enacted in 1965 by President Lyndon Johnson, prohibits federal contractors — i.e., firms with federal government contracts — from discriminating against applicants and employees based on race, color, religion, sex, sexual orientation, gender identity, or national origin.[18] The affirmative action requirements of the policy apply to all federal contractors with at least 50 employees and a contract value that exceeds (or can be reasonably expected to exceed) $50,000 in a fiscal year. All such firms must develop, implement, and annually update an affirmative action plan for each of their establishments,[19] even if the establishment is not involved in carrying out the contract.

[16] We illustrate this in Appendix Table E.2. Employment law consulting firms suggest possible discrepancies between an establishment's actual and perceived federal contractor status as well. See, for example, https://ogletree.com/insights-resources/blog-posts/i-think-my-company-is-a-federal-contractor-and-has-regulatory-obligations-but-where-can-i-look-to-search-for-that-information/, or https://www.littler.com/files/wp_covered_fed_gov_contractors_8-3-15_julie_2.pdf, or https://hortonpllc.com/what-is-executive-order-11246/. The first to point out potential issues with the accuracy of EEO-1 reports was Smith and Welch (1984), who documented large discrepancies between employment patterns by occupation in the CPS vis-à-vis the EEO-1 reports.

[17] Two papers have previously studied impacts on earnings. Holzer and Neumark (1999) studied effects of the regulation on earnings for the years 1992–1994, using survey data on 800 establishments. Leonard (1984) estimates the impact on earnings using the relative earnings growth of Black workers and the size of the federal contractor sector.

[18] Initially, this employment-based affirmative action law did not include protections for women; this was amended in 1967 to incorporate "sex" and further expanded in 2014 by President Barack Obama to include sexual orientation and gender identity.

[19] This requirement was introduced in 1968 by the enforcement agency, the OFCCP.

The affirmative action plan (henceforth AAP) has two main components. First, firms must identify jobs in which they are employing protected groups[20] at rates lower than their share in the local and qualified labor force. Second, they must develop policies to rectify any underrepresentation.

Federal contracts are awarded by auction and managed by the requesting government agency. The firm's workforce composition does not affect whether they are awarded a contract, except for very large (those valued above $10 million) or indefinite-delivery contracts, in which firms may undergo pre-award checks that include an assessment of their workforce diversity. Firms with yearly federal contracts under $50,000 and those with fewer than 50 employees are not required to develop an affirmative action plan. However, all federal contractors with contracts of $10,000 or more are subject to non-discrimination provisions from EO 11246. They are required to include an Equal Employment Opportunity (EEO) non-discrimination tagline in job postings, put up EEO posters on their websites and work sites, and maintain employment records. We do not consider these smaller firms to be regulated by affirmative action in our main analysis, because they do not develop affirmative action plans or undergo audits. This is in keeping with the literature. We will show that our results hold if we study the onset of any federal contract instead of those that pass the contract and employment size requirement for the affirmative action part of the policy.

3.1 Developing an Affirmative Action Plan

The first component of the AAP involves identifying underrepresentation. This is done by breaking down employees within the organization by specific job titles, along with their race[21] or gender. For instance, this might involve identifying the number of Black general managers or female clerical workers. They then aggregate the total employment by race and gender within broader job categories. For businesses with fewer than 150 employees, they may use broad EEO-1 job group categories;[22] larger firms must consider more detailed groups. The

[20]These groups include individuals who have been historically discriminated against or overlooked due to race, ethnicity, gender, or sexual orientation, and particularly Blacks, Hispanics, women, and non-whites. We study each of these groups separately but the paper focuses on Black workers following Miller (2017) since our findings are similar across all groups, as discussed in Section 5.

[21]For consistency and simplicity, we follow the literature and do not distinguish race from ethnicity. Most studies analyzing EO 11246 have adopted this definition due to data limitations, since EEO-1 records classify Hispanic workers as a distinct, nonoverlapping group. We adopt this classification and refer to race broadly, including Hispanics, whom we define as non-Black individuals with Hispanic ethnicity.

[22]The groups are: officials & managers, professionals, sales workers, office and clerical, craftworkers, operatives, and laborers. See https://www.dol.gov/agencies/ofccp/manual/fccm/1f-review-executive-order-aap-and-itemized-listing-data-acceptability-0.

firm then conducts an "utilization analysis" by comparing these figures to local averages for qualified individuals. Firms have the flexibility to define the local qualified population for each job category based on their geographic recruitment practices and the level of specialization required for the job. For example, although a firm may conduct a nationwide search for a CEO, they may choose to recruit for other positions within the state where they are located.

The US Census Bureau in consultation with the EEO-Tabulation Consortium provides suggested benchmark numbers for each group of the local qualified population. They use 5-year American Community Survey (ACS) data to compute the proportions of workers by race, ethnicity, and sex categorized by geography, industry, and occupation.[23] However, federal contractors are not mandated to use these benchmarks and have the flexibility to measure local qualified employment levels based on different occupational groupings and geographies, or internal firm numbers of promotable employees, as long as they can justify their choices. Regardless of how they choose to compute these reference numbers, firms are required to set a target if the proportion of workers in a job group-race-gender cell is less than 80% of its rate in the local qualified population.[24] They are then required to develop and implement policies to achieve the target. Crucially, it is illegal to use quotas or make employment decisions explicitly based on race or gender. Instead, affirmative action takes the form of policies that attract qualified candidates or improve screening mechanisms. For example, such policies could involve reaching out to professional recruitment sources and universities with a significant proportion of minority students to solicit qualified applicants from minority groups, conducting exit interviews to understand the reasons for high turnover among certain groups, or checking automated résumé screening systems for bias.[25]

3.2 Enforcement and Compliance

The Office of Federal Contract Compliance Programs (OFCCP), a division of the U.S. Department of Labor, is responsible for enforcing the policy through audits. Each year, the OFCCP audits about 1% of federal contractors. The agency may focus on certain industries, sectors, geographic regions, or types of employment practices. However, it does not specifically target individual firms or establishments unless there are employee complaints, which trigger com-

[23] These benchmarks are accessible at: https://www.census.gov/topics/employment/equal-employment-opportunity-tabulation/guidance.html.

[24] This target will be equal to 80% of the local qualified level.

[25] See the sample AAP available at https://www.dol.gov/sites/dolgov/files/ofccp/regs/compliance/AAPs/Sample_EO11246_AAP_final_01.03.18_Contr508.pdf. The example regarding screening software for biases is based on discussions with affirmative action consultants.

plaint investigations that are separate from audits.[26]

The audit process can vary based on the agency's budget[27] and priorities. Generally, it involves a desk audit followed by an on-site visit. During the audit, an OFCCP agent reviews the AAP to ensure that policies were set where needed and verify their implementation. They also examine firm applicant and pay records for any signs of discrimination, and check for the presence of EEO posters and job adverts' taglines.

Around 82% of audits result in a notice that the firm is in compliance. In case of noncompliance, the agency's responses range from seeking conciliation agreements to financial penalties or even barring the contractor from federal contracts. In practice, conciliatory agreements are the most common penalty (representing 91% of noncompliance cases). This is an agreement between the OFCCP and the contractor that outlines remedial actions that the latter agrees to implement. These agreements may also include monetary and other forms of relief, such as back pay, job offers, promotions, and retroactive seniority designed to compensate victims of employment discrimination (GAO, 2016). From 2004 to 2020, the OFCCP imposed financial penalties in less than 2% of noncompliant cases, and debarment from federal contracts, the most severe penalty, has been rare.[28]

4 Data

We build our dataset by linking information on federal contractors and audits with administrative matched employer-employee data. Below, we briefly describe each of our datasets, then explain how we construct our sample and key variables.

[26] See https://www.dol.gov/sites/dolgov/files/ofccp/CAGuides/files/WhatFederalContractorsCanExpect-CONTR508c.pdf.

[27] See Appendix Table G.1.

[28] A GAO (2016) report finds that debarment has occurred fewer than once a year since 2010. Our audit data do not include this information, but the System for Award Management, which covers active exclusions from 2015 onward, shows only 10 establishments belonging to 4 firms debarred by the OFCCP. These exclusions took place during the fiscal years 2013 and 2014 — a period that saw 7,904 audits, which highlights the rarity of such penalties. Moreover, only one of these establishments continues to be excluded from receiving federal government contracts.

4.1 Data Sources

4.1.1 Federal Procurement Data System (FPDS)

The FPDS dataset[29] contains records for federal contracts worth at least $2,500 from 2001 onward. The data include detailed information on both the government agency awarding the contract and the recipient establishment. For each establishment, we observe the firm name and address, as well as the timing and value of all federal contracts it receives. We use this information, along with firm size from census data, to determine when an establishment becomes regulated by affirmative action.

4.1.2 Audit Records from the OFCCP

Audit records from the OFCCP contain audits that were conducted each year from 2004 onward.[30] For each audit, we observe the name of the audited establishment and its parent company, as well as the audit's outcome. In most cases, the firm is found to be compliant or issued a conciliatory agreement with no financial penalties, though we also observe instances of financial retributions. The records state whether the audit identified technical infractions or evidence of discrimination. Technical violations include failing to maintain records or an affirmative action plan, while evidence of discrimination encompasses issues in recruitment, hiring, salary, accommodations, selection or testing, termination, and promotion. Combined with our administrative data, the audit records tell us which regulated establishments were audited, when, and the audits' results.

4.1.3 Longitudinal Employer-Household Dynamics (LEHD)

Our establishment outcomes are computed from the LEHD dataset, which is a linked employer-employee dataset constructed from unemployment records. We have access to these data for 18 states and Washington, D.C.[31] from 2001 to 2014. To assemble the data, we follow Sorkin (2018) and Vilhuber et al. (2018).[32] The dataset comprises workers' quarterly employment

[29] The FPDS is publicly available and can be downloaded at https://usaspending.gov.

[30] We acquired the data through a FOIA request. While we have data from 2001 onward, the coverage is lower before 2004.

[31] We have access to the following 18 states and Washington, D.C. that approved our research proposal: Arizona, Arkansas, Colorado, Delaware, Hawaii, Indiana, Iowa, Kansas, Maine, Maryland, Montana, Nevada, New Mexico, Oklahoma, Pennsylvania, Tennessee, Washington, and Wisconsin, as shown in Appendix Figure A.1.

[32] We base our data construction codes on Sorkin's (2018), which are available at https://drive.google.com/file/d/1nbreNXXKAA1cjs70-fWYGMMtZkRiJsSe/view.

histories, which we aggregate to the yearly level. We observe worker characteristics, including race, gender, and education, as well as indicators for whether these characteristics have been imputed. We use these demographic characteristics to construct establishment outcomes, such as the share of Black workers and the retention rate, as described in Section 4.3.

On the employer side, we observe characteristics such as employer size, age, metropolitan statistical area (MSA), and industry. The data include the quarterly earnings for all employer-employee pairs. The LEHD identifies employers by a Census classification SEIN code, as well as by EIN.

4.1.4 The Longitudinal Business Database (LBD) and the Standard Statistical Establishment List (SSEL)

We cannot directly link the FPDS data and audit records with the LEHD, because they have no common identifiers. Furthermore, EO 11246 applies to the entire firm, based on firm ownership, which may be broader than indicated by the SEIN, EIN, or employer name. To address this gap, we use a broader definition of the firm from the LBD, which provides a firm identifier known as *alpha*. This identifier is broader than the EIN or SEIN and helps capture the full scope of firm ownership, including conglomerates.[33] We can tackle both the lack of common identifiers and the potentially limited scope of SEIN or EIN identifiers by linking all our datasets through the LBD and SSEL. Establishments in the SSEL can be matched with their counterparts in the FPDS and OFCCP audit data on name and address. The SSEL can be linked to the LBD by a common establishment identifier, and to the LEHD by the EIN and county.

4.2 Identifying the first year an establishment is regulated or audited

Our main strategy for studying the effect of affirmative action is an event study design that accounts for dynamic effects, in which the event is first becoming subject to affirmative action. To define the treatment indicator, we leverage detailed information on contract acquisition from the FPDS and establishment-level characteristics from the LEHD. We match these datasets using a fuzzy matching method based on the term frequency-inverse document frequency (tf-idf) algorithm, which we modified to consider two dimensions: establishment name and address. The details of this procedure are discussed in Appendix B. While a fuzzy matching process may not capture all potential matches, any unmatched federal contractors are omitted for all

[33] Since some conglomerates may span multiple industries for which the variable alpha may be *too* broad, we run a robustness check that restricts our sample to cases in which defining regulation based on the variable alpha or the EIN does not affect the regulation status.

years. Assuming that any missed matches are not systematically linked to their response to affirmative action policy, this exclusion does not skew our results. We believe this assumption is reasonable, because the likelihood of missing a match depends solely on how the firm's name and address are recorded.

To identify when an establishment was first regulated by affirmative action, we look for the first year it was part of a firm — as defined by the broad firm identifier alpha — that received a contract over $50,000.[34] In any given year, all establishments in a firm will be regulated if *any* establishment has such a contract. However, we define the timing of the first regulation at the *establishment* level, as establishments may be new or may have previously belonged to a different firm.[35] For example, if a firm receives a federal contract in 2004 and merges with a (previously unregulated) firm in 2005, the resulting firm will have some establishments first regulated in 2004 and some in 2005.

Similarly, we determine the timing of the first observed audit — if any — by fuzzy matching on the establishment's name. In this case, we cannot pinpoint which establishment of a firm was audited, although we do know its state. Our main specification considers all establishments to be audited in a year if any establishment in a firm is audited. As an alternate definition, we consider a firm to be audited if any of its establishments in the same state is audited. Just as with defining the first year of regulation, the first time an establishment is audited may vary within a firm based on new establishments or mergers.

We drop establishments that were first regulated prior to 2004 to create a pre-period. We also drop banks from our sample, because they may be regulated for reasons other than holding contracts.[36]

4.3 Measuring establishment-level workforce outcomes

To measure establishment outcomes, we merge our data with the LEHD. The only common identifier between the LEHD and the LBD is the EIN, so we cannot link exact establishments between the LEHD and LBD. Instead, we merge the data on EIN, county, and year. This means that from here on, we define an establishment as an EIN-county cell. Therefore, our definition

[34] Note that we are looking for a single contract over $50,000, not contracts adding up to $50,000.

[35] It is important to note here that the firm identifier alpha may not be constant across time. We get around this by using alpha within a year and the establishment variable *lbdnum* to link establishments across years.

[36] A bank is considered a federal contractor under Executive Order 11246 if it has a bona fide federal contract, is a fund depository (typically involving Treasury Tax and Loan accounts), or processes U.S. savings bonds. Before 2015, almost all banks were subject to affirmative action because they processed savings bonds. However, the processing of savings bonds is now conducted electronically through the Federal Reserve Board, changing the affirmative action laws they are subject to after our sample period (2001–2014).

of "establishment" may encompass multiple locations of a firm within a county. When the year of first regulation (and first audit) differs across these locations, we use the modal value. We show that our results are robust to restricting our sample to observations in which this is unambiguous. Fortunately, we could use the precise identifiers of firm and establishment from the LBD to determine when each establishment is first regulated, which is where having these precise identifiers is essential for obtaining unbiased results. We consider four main outcomes. First, we measure the share of Black workers in an establishment as the number of unique Black workers over a year divided by the total number of unique workers. Since we construct this variable at the establishment level, a worker with multiple jobs over the year will be counted under each of them.[37] Next, we measure the share of Black new hires in an establishment as the number of Black new hires divided by the number of total new hires, and count an employee as a new hire if they returned to a previous employer after a gap of at least one year. We measure earnings for Black workers as the average annual earnings of all Black workers in an establishment. Finally, we measure the retention rate for Black workers as the proportion of unique workers who are observed in the firm at some point in the following year. For supplemental analysis, we construct analogous outcomes for other racial groups and for women.

Since affirmative action applies to firms with at least 50 employees, we restrict our sample to firms with at least 50 employees in *each* year they are observed, regardless of whether they have a federal contract that year. We do so to avoid comparing firms with more or fewer than 50 employees because firms face additional regulations such as the Family and Medical Leave Act — which requires that employers provide job-protected unpaid leave — and many provisions of the Affordable Care Act — all of which apply to firms with 50 or more employees — which could bias our results. Furthermore, we prefer to identify the time of first regulation only based on the first time a firm received a big enough federal contract rather than the first time a firm with a contract crosses 50 workers. This is because a change in employment may be correlated with changes in workforce composition, which could lead to reverse causality whereby trends in our outcomes determine the timing of regulation.

[37] This approach might raise concern about the differential time workers spend in different jobs, since our measure counts workers with multiple jobs within a year at each establishment, which could skew our measure of the time workers spend at each job. However, we also analyze the retention rate to account for employment stability and ensure that our results are not influenced by short-term job changes.

5 The Effect of Employment-Based Affirmative Action

We study the efficacy of affirmative action policy using an event study design. The 'event' here is the year an establishment is first observed to be subject to affirmative action. To exclude firms that are always regulated and those that frequently move in and out of regulation, we drop establishments that we observe to be regulated at any point from 2001 to 2003. While this does not guarantee that our treated establishments have never been regulated in the past, it ensures that they have not been subject to the regulation for at least 3 years.[38]

We start by studying the evolution of the establishment's proportion of Black workers around the time of regulation.[39] We then analyze the inflows and outflows of employees into establishments by examining the share of Black workers among new hires and their relative retention rate, respectively.[40] EO 11246 also stipulates good faith efforts to improve diversity through promotions. However, since we do not directly observe promotions, we instead look at (log) yearly earnings, which we expect to rise if Black workers are promoted at higher rates.[41] While we focus our analysis on Black workers, we also study each of these outcomes for Hispanics, all racial minorities, and women. Our results are qualitatively analogous across all groups.

We study the impact of becoming subject to affirmative action up to 5 years before and after the time of treatment. We control for firm age and establishment size with linear and squared terms and include year fixed effects. In our main specification, we weight establishments by size and cluster standard errors at the firm level. Given the staggered timing of the regulation and to allow for dynamic effects of the policy both over time and across treatment timings (see Goodman-Bacon (2021); De Chaisemartin and d'Haultfoeuille (2020); Sun and Abraham (2021)), we estimate our model with the estimator proposed by Callaway and Sant'Anna (2021).

We compare treated establishments with a control group that consists of establishments that become federal contractors at some point in our sample but do not receive a contract over

[38] Miller (2017) uses a pre-period of 6 years. However, as we show in Section 5.3, the discrepancy between our findings and his are independent of this difference.

[39] Employment share of protected workers is the outcome most frequently studied in previous work (Miller, 2017; Kurtulus, 2016, 2012; Rodgers and Spriggs, 1996; Leonard, 1984; Smith and Welch, 1984; Heckman and Wolpin, 1976; Goldstein and Smith, 1976; Ashenfelter and Heckman, 1974; Burman, 1973).

[40] Previously only studied by Holzer and Neumark (1999, 2000a) with a survey the authors conducted covering 800 establishments in Atlanta, Boston, Detroit, and Los Angeles between 1992 and 1994.

[41] As we will show, we find no sizeable effects on retention. If we did, however, this would be a biased measure of wages because we do not observe whether workers spent the full quarter in the establishment, but we do not find such effects.

$50,000 in any year and are therefore not subject to affirmative action.[42] Like the treated establishments, we require that all establishments in the control group belong to a firm with at least 50 employees in all years.[43]

We estimate the parameters:

$$\beta_{AA}^{Y_t}(g,t) = \mathbb{E}\big[Y_t(1) - Y_t(0) | X_j, \lambda_t, \mathbb{1}(j \text{ is first regulated at time g}) = 1\big], \tag{1}$$

where $Y_t(1)$ and $Y_t(0)$ denote the potential outcome at time t if an establishment is regulated and unregulated, respectively. So, $\beta_{AA}^{Y_t}(g,t)$ captures the effect of affirmative action on outcome Y_t for units treated at time g conditional on firm characteristics at time $g-1$, X_j,[44] and year fixed effects, λ_t. We identify the parameters of interest by comparing the outcomes of treated units with those of never-treated ones, where the never-treated units consist of the set of establishments with federal contracts under $50,000. We cluster standard errors at the firm level; the level at which the policy is defined.

The parameters in equation (1) allow for heterogeneous effects across treatment groups and over time. To interpret them, we aggregate these group-time effects into the event study parameters $\{\theta_{AA}(\tau)\}_{\tau=-4}^{5}$ capturing the effect of being exposed to affirmative action in employment for exactly $\tau = t - g$ years.[45]

Our identification assumption is that the outcome variables are affected by the timing of regulation through, and only through, the onset of affirmative action. Specifically, the time at which an establishment receives a large enough contract — and thus becomes subject to affirmative action — depends on the availability and profitability of government contracts rather than on the compliance costs of affirmative action. This assumption will be violated if firms time their federal contracts according to their current workforce composition. We discuss how we deal with potential threats to identification in Section 5.2.

[42] An alternative control group would consist of "not yet treated" establishments. However, this would limit the number of lags we can estimate and require that we classify soon-to-be-treated establishments as controls.

[43] Given the myriad of policies that apply to firms with 50 or more employees, firms above and below this threshold are not comparable.

[44] We include pre-treatment covariates to ensure that the assumption of conditional parallel trends holds as can be seen by the pre-treatment estimates being statistically indistinguishable from zero.

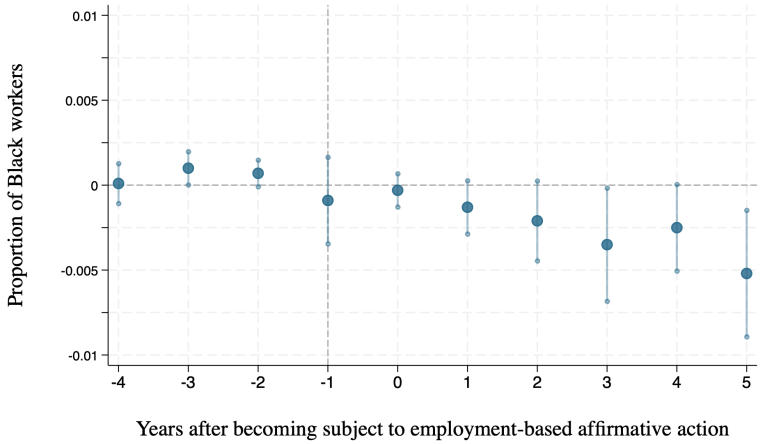
[45] These are the parameters we interpret throughout and are weighted averages of $\{\beta_{AA}^{Y_t}(g,t)\}_{t,g}$ given by

$$\theta_{AA}(\tau) = \sum_{g=2}^{T} \mathbb{1}(g+\tau \leq T)\beta_{AA}^{Y_{g+\tau}}(g,g+\tau)\mathbb{P}(G=g|X_j,\lambda_t,G+\tau \leq T). \tag{2}$$

See Callaway and Sant'Anna (2021), Section 3.1.1. for details.

5.1 Results

Figure 1: *Proportion of Black workers upon becoming subject to affirmative action*



Notes: The figure plots our estimated dynamic effects of affirmative action on the establishments' share of Black workers, measured *tau* years after exposure to the policy. We cluster standard errors at firm level. Bars indicate 95% confidence intervals. The sample size rounded to the nearest 1,000 used for estimation is 750,000. The Wald test for joint significance of the post-treatment coefficients, those corresponding to $\tau = 1, \ldots, 5$, has a p-value of 0.0015. Therefore, we can reject the null of the joint effect of such coefficients being zero at the 1% confidence level.

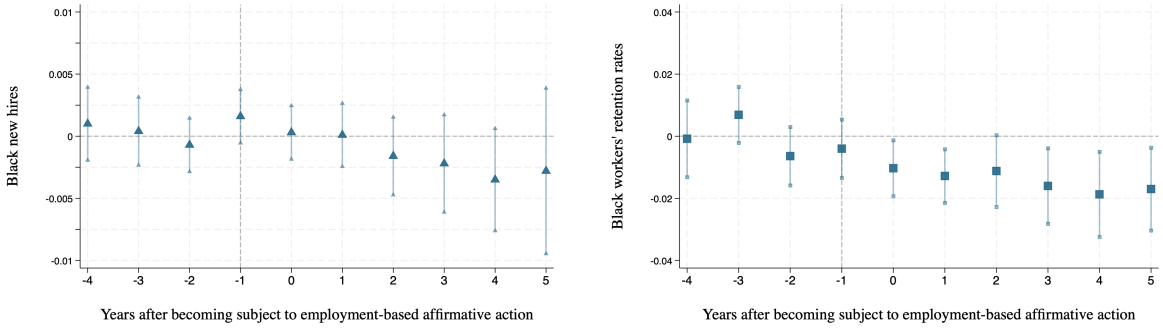
Figure 1 shows no evidence of a positive effect of affirmative action on the share of Black workers in an establishment. Our results are mostly statistically insignificant and quantitatively small; if anything, the estimates post-treatment trend negative. The only statistically significant effect we find is 5 years after treatment, showing a decrease of half a percentage point in the share of Black workers, whose baseline value is 0.12. While the effects are jointly significant — the p-value of the Wald test for joint significance is 0.0015, — these effects contrast with those found by Miller (2017). We discuss the differences and show that the discrepancy is explained almost fully by the different periods studied in Section 5.3. A possible explanation for the small effects we find could be that minorities are hesitant to work at regulated firms due to concerns of tokenism (Leibbrandt and List, 2018). We come back to this point after discussing the effects we find in employment flows.

Since an establishment may take time to adjust the workforce, we might expect to see an increase in the share of Black hires before the overall stock of workers significantly changes. However, we observe no positive effect on the share of Black new hires. For reference, the baseline share of Black new hires before becoming contractors is 13%, so the effects in Figure 2 (a) are relatively smaller than those in Figure 1. Additionally, a Wald test of joint significance

cannot reject the null hypothesis of all coefficients being zero (the p-value is 0.4522). Another potential channel for workforce adjustment is retention. But as seen in Figure 2 (b), we find no evidence of positive effects for Black workers here either. In fact, Black workers are *less* likely to remain in treated firms; this result is small in magnitude but statistically significant. The estimated effects represent, on average, a 2.3% decrease from the pre-treatment retention rate of 64%.

Figure 2: *Inflows and outflows upon becoming subject to affirmative action*

(a) *Black workers' share of new hires* ***(b)*** *Black workers' retention rates*

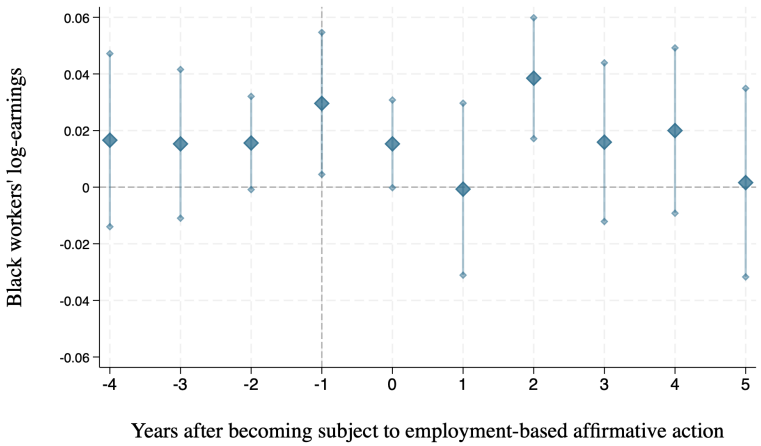


Notes: The figure plots our estimated effects of becoming subject to affirmative action in employment on the inflows and outflows of Black workers at establishment level τ years after exposure. ***Panel (a)*** shows the estimated effects of the treatment on the establishments' share of Black new hires, measured τ years after exposure to the policy. The sample size rounded to the nearest 1,000 used for estimation is 653,000. The Wald test for joint significance of the post-treatment coefficients, those corresponding to $\tau = 1, \ldots, 5$, has a p-value of 0.4522. Therefore, we cannot reject the null of the joint effect of such coefficients being zero. ***Panel (b)*** shows estimated effects of the treatment on the retention rates of Black workers in treated establishments. The sample size rounded to the nearest 1,000 used for estimation is 450,000. The analogous Wald test for the effects on retention rates post-treatment has a p-value of 0.0000, so we reject the null of these effects being jointly zero at the 0.1% level. Bars indicate 95% confidence intervals. We cluster standard errors at firm level.

Taken together, the null or small negative effects we find on the share of Black workers and on the inflows and outflows of these workers at regulated firms suggest that affirmative action in employment does not trigger sizeable changes in workforce composition. A possible explanation for our results could be that minorities hesitate to work in regulated firms due to concerns of being seen as symbolic hires rather than being valued for their qualifications. Leibbrandt and List (2018) investigate this phenomenon by analyzing the impact of Equal Employment Opportunity (EEO) statements in job advertisements on the application rates of minorities in a large-scale natural field experiment across 10 cities in the US. Their findings reveal that EEO statements actually deterred racial minorities from applying due to perceptions of "tokenism." That said, we hesitate to interpret the negative results, because they are mostly insignificant or quantitatively small.

Next, since affirmative action stipulates efforts to increase diversity through promotions, we turn to study the impact of affirmative action on the yearly earnings of workers as we do not directly observe promotions: Black workers' average yearly earnings, would rise if they were promoted more frequently. But again, we do not see significant changes. While we see higher earnings amongst firms receiving a federal contract, we see no changes over time, and the magnitudes of the effects are quite small: even considering the largest effects within the 95% confidence bands, they represent at most 0.06 log points of the benchmark level.

Figure 3: *Black workers' log earnings upon becoming subject to affirmative action*



Notes: The figure plots our estimated effects of affirmative action on the establishments' log earnings among Black workers, measured τ years after exposure to the policy. We cluster standard errors at firm level. Bars indicate 95% confidence intervals. The sample size rounded to the nearest 1,000 used for estimation is 309,000. The Wald test for joint significance of the post-treatment coefficients, those corresponding to $\tau = 1, \ldots, 5$, has a p-value of 0.0037. Therefore, we reject the null of the joint effect of such coefficients being zero at the 1% confidence level.

These results are not particular to Black workers; we also find no positive impact on Hispanics, all racial minorities as a group, or women.[46]

5.2 Identification and Robustness

We next address potential sources of bias that could cause us to miss any true positive effect. First, we may be concerned that becoming a federal contractor affects firm composition and wages through channels other than affirmative action. For example, a federal contract may increase the establishment's labor demand, which may differentially affect Black workers. How-

[46]Due to the Census Bureau's disclosure avoidance requirements, this version of the paper withholds some numerical results and summary statistics, including these. This approach allows us to preserve full flexibility for the final version of the paper.

ever, Aizer et al. (2020) document — in the context of World War II — that an increase in labor demand due to government contracts tends to *increase* minority hiring, so this force would likely amplify rather dampen our estimates. At least to an extent, the control group of federal contractors below the $50,000 threshold mitigates this concern. Additionally, we supplement our findings with a regression discontinuity design that compares current contractor firms above and below this threshold to separate the impact of the regulation from other aspects of receiving a federal contract. Our sample for the RD consists of federal contractors in the year they first receive a contract within our defined bandwidth, which covers contract values between $25,000 and $75,000. We exclude any contractors that previously held a contract over $75,000. Using this design, we again find no significant jump in the share of Black workers, or the share of Black hires, across the threshold in the year they became federal contractors or 1 year later. Since this design does not rely on time variation upon becoming a federal contractor, it supports that our findings are not driven by any other effect of becoming a federal contractor that might cancel out effects from the policy. We provide further details and validity checks of the RD in Appendix C.

Another concern is that firms may be entering our sample already regulated, either through sub-contracts or state-level policies. Importantly, indirect regulation would attenuate but not reverse the generally negative (though mostly insignificant) effects we find. Furthermore, firms entering our sample as 'indirectly' regulated, would only become subject to enforcement audits after entering our sample, as subcontractors and employers complying with state-level policies are not subject to these audits. Additionally, we would expect to see impacts on establishments that start with minority employment shares below reference rates, and we will show, this is not what we find.[47] We find no significant effects even in this subsample, in support of the robustness of our results. The results from our regression discontinuity design further corroborate our results by comparing firms just above and below the regulation threshold at a single point in time, rather than relying on changes over time.

Our analysis focuses on establishments that receive federal contracts above the $50,000 threshold, as these are subject to the full affirmative action policy we study. The control group consists of establishments with contracts below this threshold, the majority of which are only required to comply with the minimal provisions of EO 11246. These provisions include displaying Equal Employment Opportunity (EEO) posters, incorporating EEO taglines in job advertisements, and maintaining records of hiring, firing, and promotion decisions and apply to

[47]The *reference rates* we refer to here are the EEO-Tabulations which are a reference point for federal contractors to assess representation. These are not quotas.

any federal contractor with a contract value above $10,000 (see Section 3). To ensure that our results are not biased downward by potential positive effects of these minimal requirements on the control group, we conduct an event study, around when a firm first receives a federal contract with a value above $10,000. Our analysis reveals no significant effects on any of our outcome measures for this group.

Selection into federal contracting presents another concern. Firms with lower shares of minority workers may perceive their compliance cost as higher and bid for contracts under $50,000 or not bid for contracts at all. If firms may underbid due to the policy, we should expect to see that the value of federal contracts cluster right below $50,000. As seen in Appendix Figure D.1, this does not seem to be happening: if anything, there is a cluster of contracts just *above* the threshold, and pre-contract outcomes are continuous across the threshold. To the extent that some firms may still avoid federal contracting altogether due to the policy, we cannot identify the counterfactual treatment outcome for these firms, and our results represent the effect of the treatment on the treated.

Firms may also preempt the regulation in the years before becoming a federal contractor by making efforts to increase workforce diversity *before* receiving a contract, so there is no remaining effect on the workforce once they do. But in this case, we would have expected to see positive pre-trends, which we do not find.

Last, we perform a set of technical robustness checks. Our results are robust to modifying the specification by weighting all establishments equally or not clustering standard errors. Since our definition of an establishment here is an EIN-county cell, it is possible that this group includes locations that were first regulated at different times.[48] Restricting our sample to establishments for which the first year of regulation is unambiguous does not change our results. In the case of very large conglomerates that span industries, it is a gray area whether the entire conglomerate counts as a firm for the regulation purposes. To analyze the effect of the policy in these cases, we construct an alternate measure of regulation, assuming that the "firm" is the more restrictive EIN, instead of the firm indicator alpha from the LBD. Our results are robust to only considering establishments for which these two measures coincide. Finally, removing observations in which race and ethnicity are imputed[49] does not change the results.

[48] In a given year, regulation is always at the firm level. However, some establishments may not have been part of the firm at that time.

[49] 75% of LEHD records have race and ethnicity attached using a direct Social Security Number link, but 20% of the records have imputed race and ethnicity information from the American Community Survey. Only 5% of records have no person-specific information on race and ethnicity.

5.3 Benchmarking our findings to Miller (2017)

Miller (2017) conducts a similar event study analysis and finds positive, sizeable, and persistent effects of affirmative action on the proportion of Black workers from 1978 to 2004. We show here that the discrepancy in our results is driven by the time frame rather than by differences in the data between our studies. We replicate Miller's (2017) analysis using his data source,[50] and replication codes,[51] changing *only* the period to match ours (2001–2014).

Figure 4(a) shows our event study estimates using EEO-1 reports in our sample period (2001–2014) together with the estimates we obtain with our novel data. We overlay an approximation of Miller's (2017) estimates (see Panel A of Figure 3) to illustrate the difference between Miller's (2017) estimates using EEO-1 reports for 1978-2004 and ours using both EEO-1 reports and our data for 2001–2014.

Next, to illustrate that the difference in timing largely explains the difference in our results with respect to Miller's (2017), we split the EEO-1 reports we have available into twelve overlapping 14-year periods, with our sample period (2001–2014) being the most recent one. Figure 4(b) shows the estimates for each of these periods, starting with 1990–2003 showing the largest effects, and ending with our sample period showing effects closest to zero. We overlay an approximation of Miller's (2017) estimates (see Panel A of Figure 3) for context. The comparison reveals that the regulation's effectiveness remained strong until as recently as 1990–2003 but declined steadily thereafter.

Miller (2017) finds that 5 years after first becoming a contractor, the establishment Black share increases by about 0.88 p.p. Our results for 1990–2003 show a similar increase of about 0.90 p.p. However, for our period (2001–2014), the increase is only 0.10 p.p. and none of the point estimates are significant at the 95% confidence level.

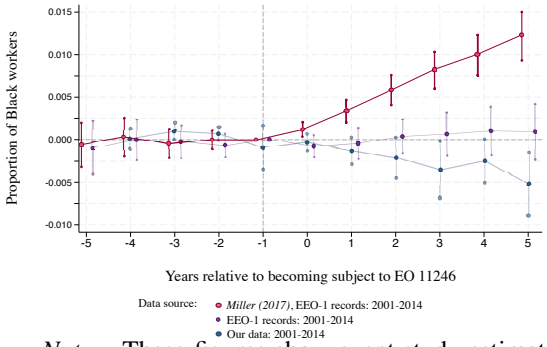
To the extent that results from the EEO-1 reports trend positive (though statistically insignificant and negligible in magnitude) compared with our main results, this discrepancy may stem from how regulation status is measured: EEO-1 reports, filed annually by employers with at least 100 employees and federal contractors with at least 50 employees, require self-reported

[50] We got access to EEO-1 reports, the data Miller (2017) uses, for the period 1990–2014, which has 14 years of overlap with his sample period (1978—2004).

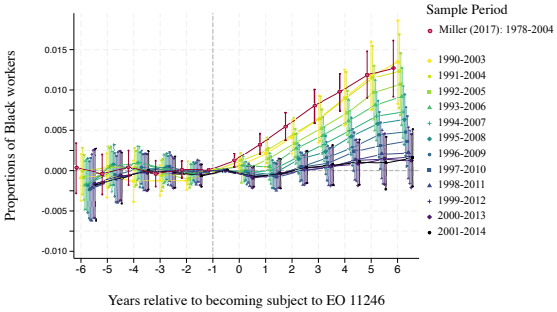
[51] Specifically, we replicate his *regulation event study* and estimate $y_{jt} = \alpha_j + \lambda_{d(j),t} + \gamma X_{jt} + \sum_{k=-5}^{k=5} \theta_k D_{jt}^k + \varepsilon_{jt}$ using two-way fixed effects (TWFE), where j indexes establishments and t indexes time. α_j represents establishment fixed effects, $\lambda_{d(j),t}$ represents MSA-by-year fixed effects, and X_{jt} is a quadratic polynomial of the log of establishment size. D_{jt}^k are leads and lags of establishment j first becoming a contractor, defined as $D_{jt}^k = D_j \mathbb{1}(t = \tau_j + k)$, where D_j is an indicator for establishment j first becoming a contractor, and τ_j is the year the establishment first becomes a contractor. The control group consists of non-contracted establishments throughout the period.

Figure 4: *EEO-1 Reports: Share of Black Workers Upon Becoming Subject to EO 11246*

(a) *Event Study Estimates Over Time*



(b) *Comparison with Miller (2017)*



Notes: These figures show event study estimates using EEO-1 reports from 1990 to 2014. ***Panel (a)*** shows event study estimates using EEO-1 reports in our sample period (2001–2014) together with the estimates we obtain with our novel data. We overlay an *approximation* of Miller's (2017) estimates for the period 1978–2004 as a benchmark. Bars represent 95% confidence intervals. We normalize $\theta_{-1} = 0$ and cluster standard errors at firm level. ***Panel (b)*** displays event study estimates of the leads and lags for an establishment first becoming subject to affirmative action and covers 12 time periods. The earliest period starts in 1990 and ends in 2003, and the latest corresponds to our study period with novel data from 2001 to 2014. Each time window spans 14 years, but the *first* time an establishment is defined to be under affirmative action, it is restricted to not having been regulated since 1990, the first available year.

information about affirmative action regulation and workforce diversity. Each establishment of the firm with at least 50 employees is required to file a separate report. Notably, approximately 6% of firms report inconsistent 'affirmative action statuses' across their establishments, despite the regulation applying at the firm level.[52] Among these firms with mixed reporting, on average 45% of their establishments report not meeting the criteria for affirmative action when other units report the opposite.[53] Furthermore, since only establishments meeting the size thresholds are required to submit EEO-1 reports, it is possible that some eligible establishments fail to report due to a lack of awareness of their regulatory obligations. This potential underreporting could mirror the misreporting observed in submitted forms.[54]

[52] The specific question that EEO-filers answer regarding 'affirmative action status' is whether the company or any of its establishments have 50 or more employees and have a federal contract (or subcontract) of $50,000 or more.

[53] See Appendix Table E.2.

[54] Numerous posts from consulting firms and employment law firms suggest possible discrepancies between an establishment's actual and perceived federal contractor status. See, for example, https://ogletree.com/insights-resources/blog-posts/i-think-my-company-is-a-federal-contractor-and-has-regulatory-obligations-but-where-can-i-look-to-search-for-that-information/, or https://www.littler.com/files/wp_covered_fed_gov_contractors_8-3-15_julie_2.pdf, or https://hortonpllc.com/what-is-executive-order-11246/.

5.4 Are newly regulated firms already in compliance?

We now show that the lack of effects from affirmative action cannot be attributed to firms already being sufficiently diverse before regulation. Federal contractors subject to affirmative action must target minority shares representative of the local qualified labor force to create the strategies in their AAPs. If they are close to matching or match the local qualified average minority share, they are not required to develop and implement strategies to address underrepresentation in their AAPs. Only when a firm's share of Black workers falls below 80% of the local qualified average must it outline strategies to reduce this discrepancy.

To provide local qualified averages to federal contractors, the Census Bureau in consultation with the EEO-Tabulation Consortium, prepares EEO-Tabulations. These tabulations provide suggested benchmark proportions of workers by race, ethnicity, sex, industry, and occupation at various local levels for federal contractors to use in constructing their AAPs. Since the LEHD does not include occupation information, we estimate analogous tabulations by county and industry across broad education levels (college and non-college) using the same data source as the Census and the EEO-Tabulation Consortium: 5-year ACS data. We merge these estimated tabulations with our dataset to define approximate reference shares for each establishment.[55]

If the regulation had any positive effects, we would expect to see evidence of these impacts among firms below the 80% threshold. However, when we repeat our analysis using the subsample of firms below this 80% threshold, we continue to find no significant effects on any of the outcomes we study (i.e. Black share of workers, retention rates, hiring rates, or average earnings). We also examine the policy's impact on the difference between a firm's proportion of Black workers and that of the local qualified labor force, again finding null effects. These results suggest that even firms with room for diversity improvement are not effectively responding to the policy's requirements.

These results are subject to the caveat that our estimates of target minority levels may differ from the levels establishments choose for their AAPs. Establishments have considerable flexibility in defining the 'local and qualified population,' as long as they can justify their choices. For example, while we consider the local labor market to be the county, establishments may recruit at the state or even national level for some jobs, which would change the relevant labor market. Furthermore, establishments can directly consider occupation, which we impute through industry and education. While establishments must set goals for each occupation group where they underrepresent minority workers, we can only estimate underrepresentation for the

[55] See more details about these tabulations and how they compare to the EEO-Tabulations in Appendix B.1.

firm as a whole. For example, we would not identify an establishment as below target levels if it discriminated in promotions so that minority workers were underrepresented in managerial jobs but over-represented in lower paid jobs. This means that our measures of under-representation will tend toward mis-categorizing establishments as having met their targets, rather than the other way around. So, despite these caveats, we believe that the establishments we identify as underrepresenting minorities would, in fact, have been required to implement strategies in their AAPs to combat underrepresentation.

6 Do Audits have any Effect?

From Section 5.4, we see that our results are not driven by federal contracting firms that are already hiring Black workers at target rates. This suggests that affirmative action should still have scope to affect workforce outcomes, and the absence of positive effects indicates issues with compliance. We turn to enforcement of the regulation through audits to explore compliance. We start by asking how audits affect outcomes, for all establishments and then for those found to not be in compliance. As described by Section 3.2, the OFCCP enforces EO 11246 by auditing approximately 1% of contractors each year and can impose penalties — most commonly, conciliatory agreements — if the establishment is not in compliance.

We identify the effect of audits using an event study analogous to our main specification in Section 5. In this case, the event is the first year we observe a firm being audited. If any establishment in a firm is audited in a given year, we consider the entire firm to be audited. Our results are robust to requiring the audited establishment to be in the same state.[56] The control group comprises all firms that we never observe being audited. We study the impact on the proportion of Black workers in the workforce and among new hires, as well as earnings and retention rates. Under the assumption that the timing of audits is plausibly random from employers' perspective, we can identify the causal impact of audits on each of our workforce outcomes.

We argue that this assumption holds. While the OFCCP may target particular industries or regions, it did not generally target individual firms during our study period.[57] A study by the Government Accountability Office in 2016 (GAO, 2016) found that the OFCCP did not target employers based on the likelihood of noncompliance and recommended that the OFCCP start doing so.[58] Furthermore, GAO (2016) finds that the regional distribution of audits was

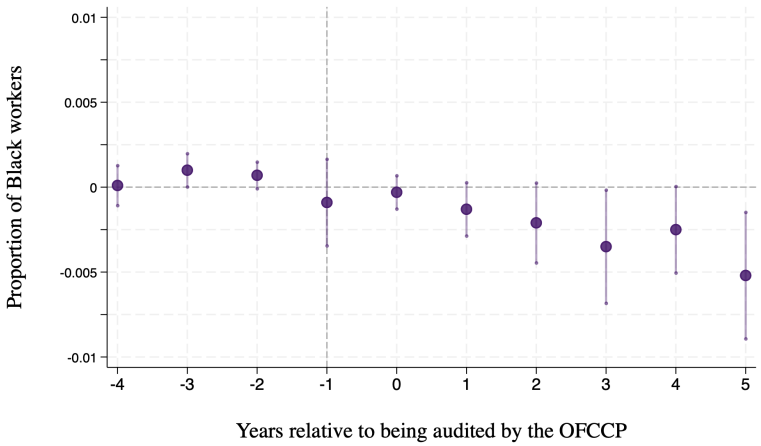
[56] We generally cannot identify the exact establishment that is audited within a state.

[57] These audits are separate from investigations in response to employee complaints, which are targeted.

[58] This change would render audits nonrandom, but it occurred after our sample period.

influenced by the regional distribution of the current workforce of the OFCCP. Even if this assumption does not hold and employers are able to predict being audited, this would bias our results downward post-audit, but any such effect would be picked up as a pre-trend, and we do not see any evidence of this.

Figure 5: *Proportion of Black workers upon being audited by the OFCCP*



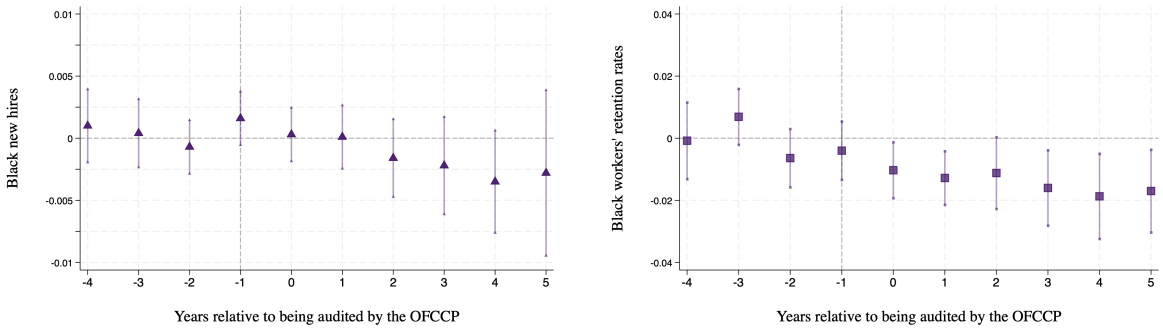
Notes: The figure plots our estimated dynamic effects of an audit for compliance with affirmative action on the establishments' share of Black workers, measured τ years after the audit. We cluster standard errors at firm level. Bars indicate 95% confidence intervals. The sample size rounded to the nearest 1,000 used for estimation is 1,714,000. The Wald test for joint significance of the post-treatment coefficients, those corresponding to $\tau = 1, \ldots, 5$, has a p-value of 0.9836. Therefore, we cannot reject that the joint effect of these coefficients is zero.

As shown in Figure 5, receiving an audit has a statistically insignificant and quantitatively small effect on the share of Black workers in an establishment's workforce or among hires. On average, the results indicate a 0.2% decrease in the share of Black workers. Similarly, Figure 6 (a) reveals that the share of Black new hires decreases by an average of 0.8%, while there are small negative effects on retention rates (Figure 6 (b)). Although these effects are jointly significant — the Wald test of joint significance has a p-value of 0.0186, — their magnitude is less than 2 percentage points from a baseline of 62%. Additionally, Figure 7 shows that the log earnings of Black workers decrease by 0.2 log points from a baseline of 9.5. Our results hold for all protected groups and are robust to the set of robustness checks in Section 5.2. These findings indicate that audits do not positively impact our outcomes of interest, which suggests that our results in Section 5.1 are not primarily driven by insufficient audits.

We next check whether audits have any effect on the subset of establishments that are not in compliance. For this, we can define 'non-compliance' in two different ways: as being found non-compliant by an audit, or underrepresenting minority workers the year before the audit.

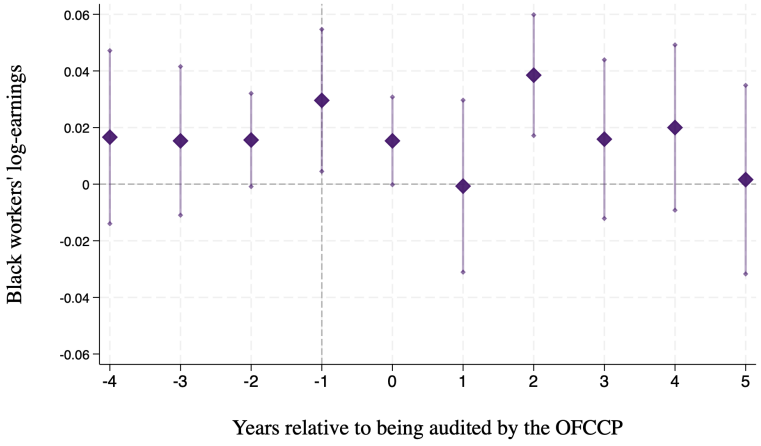
Figure 6: *Inflows and outflows upon becoming subject to affirmative action*

(a) *Black workers' share of new hires* ***(b)*** *Black workers' retention rates*



Notes: The figure plots our estimated dynamic effects of an audit for compliance with affirmative action on the inflows and outflows of Black workers at establishment level τ years after an audit by the OFCCP. ***Panel (a)*** shows the estimated dynamic effects of an audit on the establishments' share of Black new hires. The sample size, rounded to the nearest 1,000, used for estimation is 1,594,000. The Wald test for joint significance of the post-treatment coefficients, those corresponding to $\tau = 1, \ldots, 5$, has a p-value of 0.7879. Therefore, we cannot reject that the joint effect of these coefficients is zero. ***Panel (b)*** shows the estimated effects of the treatment on the retention rates of Black workers in audited establishments. The sample size, rounded to the nearest 1,000, used for estimation is 906,000. We cluster standard errors at firm level. Bars indicate 95% confidence intervals. The analogous Wald test for the effects on retention rates post-treatment has a p-value of 0.0001, so we reject the null of these effects being jointly zero at the 0.1% level.

Figure 7: *Black workers' log earnings upon being audited by the OFCCP*



Notes: The figure plots our estimated dynamic effects of an audit for compliance with affirmative action on the log earnings of the establishments' Black workers, measured τ years after the audit. We cluster standard errors at firm level. Bars indicate 95% confidence intervals. The sample size rounded to the nearest 1,000 used for estimation is 1,714,000. The Wald test for joint significance of the post-treatment coefficients, those corresponding to $\tau = 1, \ldots, 5$, has a p-value of 0.0186. Therefore, we reject the null of the joint effect of such coefficients being zero at the 5% confidence level.

We start with the first definition. Around 20% of audited establishments are not found to be in compliance. In almost all such cases, the OFCCP issues these establishments a conciliatory agreement that requires the firm to correct any issues and may include provisions such as back-pay. We find similarly null results when we restrict our sample to this group. We report these results in Appendix F. Next, we consider establishments that underrepresent minority workers. As in Section Section 5.4, we restrict our sample to the subset of establishments with shares of Black workers less than 80% of our estimates of their share in the local qualified population. Here, too, we find no impact of an audit on any of our workforce outcomes.

It may seem counterintuitive that establishments receiving conciliatory agreements — which require the firm to take corrective steps — still show no measurable impact. Establishments may receive a conciliatory agreement to fix a procedural infraction, or because the OFCCP found evidence of discrimination. The latter is more serious and could involve provisions such as backpay or retroactive hiring. However, the OFCCP only finds evidence of discrimination in 2–3% of audits.[59] Most infractions are technical violations, such as not maintaining an affirmative action plan or retaining records. Even if these technical issues are corrected, they may not lead to significant changes in our outcomes of interest, as we will discuss this further in the next section. Note that this is not to say that less than 3% of firms are discriminating against protected groups, as the OFCCP would need evidence to show that disparities were driven by discrimination. Kline et al. (2022) directly test discrimination in hiring through an audit study in entry-level positions and find that among the unambiguously discriminatory firms, 23% are federal contractors.

These results raise the question of whether audits accurately identify which establishments are not making good-faith efforts to combat disparities. In other words, how related are our two measures of 'non-compliance'? To test this, we compare disparities, as measured by the difference between our estimated EEO tabulation benchmark measures and the firm's share of Black workers, across firms found compliant vs non-compliant during audits. We find that audited firms, particularly those that pass audits, have on average slightly higher shares of Black workers than those in the local qualified population. However, these differences are small in magnitude compared to the variance in this computed disparity across establishments. Indeed, the median firm shows lower shares of Black workers than the benchmark level. These results are subject to the same caveats as in Section 5.4, due to the limitations in estimating the EEOC tabulation benchmarks. However, these limitations should apply equally to firms deemed com-

[59] Consistent with GAO (2016), we find that from 2010 to 2015, only 2% of cases have evidence of discrimination — see Appendix Table G.2. From 2004 to 2020, which is the period spanned by our data, this fraction is 3%.

pliant or not.

7 Discussion

Our analysis suggests that despite longstanding affirmative action policies, significant barriers to equal employment opportunities persist in the modern labor market, even among firms explicitly subject to these regulations. In Section 5.4, we show that while the regulation may face diminishing returns, it still has scope for continued impacts. This suggests that the persistent under-representation of minorities in regulated firms may indicate an issue with compliance. Lack of compliance here could mean firms not following the regulation or, more subtly, setting policies that do not effectively combat under-representation.

We first discuss whether EO 11246 became easier to comply with. While the regulation's effects are expected to diminish as labor market disparities decrease,[60] a general reduction in discrimination would not render the policy ineffective. This is because the target minority shares are benchmarked to their representation in the local labor force. As discrimination declines and minority worker shares increase across the board, the target shares for regulated firms would also rise proportionally. Thus, unless discrimination was fully eradicated, the policy would continue to promote change among regulated firms, albeit at a decreasing pace as the labor market's discrimination decreases.

Instead, affirmative action could have become easier to satisfy if regulated firms were already hiring minorities at rates representative of local qualified populations. In this case, firms would not need to set policies on their AAPs, as they would already be in compliance without changes to their employment or promotion practices; in other words, the regulation would not be expected to have an effect. This shift could occur if firms that hire minorities at representative rates become more likely to be awarded federal contracts.[61] To test this hypothesis, we compare the proportion of minority workers in regulated firms to their respective locally qualified benchmarks in the year prior to first regulation. We find that, on average, the proportion of Black workers in regulated firms is 2.9% higher than the benchmark populations.[62] However,

[60] As Leonard (1984) puts it "if one thought of the growth of nonwhite representation as following a logistic growth curve, then squeezing further gains from establishments in the upper tail of the distribution would be difficult."

[61] Since federal contracts are awarded and managed by the requesting government agency which is independent of the agencies that oversee compliance of EO 11246, this higher rate of awarded federal contracts to more diverse firms would be driven by these firms competing at higher rates for federal contracts.

[62] Note that due to Census disclosure procedures this analysis is currently based on a subset of firms that have been audited (representing approximately 1% of the total); this could be extended to the full sample during a

these differences are small relative to the variation in workforce composition within the firms (the standard deviation is 0.15), and at least 50% of regulated firms have a 2% lower proportion of Black workers than the benchmark population.[63] This is in line with Rodgers and Spriggs (1996) who find a similar pattern of positive selection in 1982–1992, and Kurtulus (2016) in 1973–2003. This pattern also aligns with the findings of Kline et al. (2022), who show that federal contractors discriminate less than other firms on average, but multiple federal contractors are amongst the most discriminatory firms in their sample of Fortune 500 firms. It also suggests that while regulated firms are more likely to be complying at the outset, the regulation should have a positive impact on workforce composition in the majority of firms. However, we find null effects when considering only those firms with proportions of Black workers below benchmark levels in Section 5.4. Together, this shows that despite evidence for positive selection into federal contracting, there should have been scope for further impacts.

This implies that there may be issues with compliance. Audits find 80% of establishments to be compliant. However, the OFCCP only audits 1% of firms, and it is not clear that this high compliance rate reflects meaningful affirmative action measures in all these firms. For example, a GAO (2016) report finds that 85% of firms receiving an audit did not submit their AAP within the allotted 30 days,[64] suggesting that many may delay creating or updating their APP until they are audited. This could mean that up to 85% of firms selected for audits may not have identified minority underrepresentation or developed strategies to address it. This lack of compliance may partially explain the null results we observe. Moreover, we do not see any positive impact of audits, irrespective of the audit outcome (see Section 6 and Appendix F), indicating that it is possible to pass an audit without implementing substantive changes to improve workforce diversity.

Anecdotal evidence suggests that the auditing process of the enforcement agency is thorough and that the average employer spends both time and resources preparing for them. The core issue here may be that audits are designed to assess compliance with the *letter* of the law, not the *spirit* of the law. We may expect to see different results based on whether firms treat the AAP as a bureaucratic document or as a firm-wide plan actively implemented by top managers. For the regulation to be effective, firms need to maintain an AAP that correctly identifies areas of under-representation and implement effective policies targeting these areas. Firms have

revision.

[63] Our approximate benchmarks group all occupations together, whereas, in practice, firms will likely find at least one job group where they need to set a target.

[64] During our period of study, the OFCCP only reviewed the AAP if it audited the establishment. We discuss changes implemented in 2022 to monitoring the AAP at the end of this section.

flexibility in identifying job groups and the relevant local labor markets. Savvy firms may make strategic choices to circumvent the need to create or implement meaningful policies in their AAPs. Where they do set policies, it does not necessarily follow that the policies will be effective. For example, if a firm under-represents minority workers due to biased screening mechanisms or a negative reputation amongst prospective employees, merely setting a policy that aims at advertising job openings to minority workers will not address the underlying disparity.

Overall, our analyses suggest a multifaceted story behind the null effect of EO 11246 on workforce outcomes: firms may not be maintaining an AAP if they believe the threat of an audit is low; writing an AAP may not identify or address the root causes of disparities; firms underrepresenting minority workers may be less likely to compete for a federal contract. This means that improving the regulation's efficacy must involve multiple fixes.

Despite these challenges, we are optimistic about recent changes implemented by the OFCCP, which may address some of our concerns and potentially reverse the downward trend in EO 11246's effectiveness that we uncover in Section 5.3. Since 2022, the OFCCP has launched a contractor portal that requires firms to certify to have an AAP prepared, irrespective of whether they are audited. Following the GAO (2016) report's recommendations, the OFCCP has increased the targeting of audits toward firms that are less likely to be compliant, and increased outreach efforts to educate and inform regulated firms. Based on anecdotal evidence, they are also increasingly discussing the effectiveness of the policies firms choose to implement during audits and asking firms to follow up with new policies if the current ones are not showing results. It remains a question for future work whether these changes will turn the tide and bring back the efficacy of EO 11246 seen in previous decades.

References

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock (2009) "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," in *Producer Dynamics: New Evidence from Micro Data*, University of Chicago Press, 149–230.

Aizer, Anna, Ryan Boone, Adriana Lleras-Muney, and Jonathan Vogel (2020) "Discrimination and Racial Disparities in Labor Market Outcomes: Evidence from WWII," NBER Working Paper Series.

Anderson, Bernard E. (1996) "The Ebb and Flow of Enforcing Executive Order 11246," *American Economic Review*, 86, 298–301.

Ashenfelter, Orley and James J. Heckman (1974) "Measuring the Effect of an Anti-Discrimination Program," NBER Working Paper Series.

Burman, George Robert (1973) "The Economics of Discrimination: The Impact of Public Policy," Ph.D. dissertation, The University of Chicago.

Callaway, Brantly and Pedro H. C. Sant'Anna (2021) "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225, 200–230.

De Chaisemartin, Clément and Xavier d'Haultfoeuille (2020) "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–2996.

Farrell, Jane (2020) "The Promise of Executive Order 11246: Equality as a Fact and Equality as a Result," *Journal of Social Justice*, 13, available at https://via.library.depaul.edu/jsj/vol13/iss2/4.

GAO (2016) "Equal Employment Opportunity—Strengthening Oversight Could Improve Federal Contractor Nondiscrimination Compliance," Report to Congressional Requesters.

Goldstein, Morris and Robert S. Smith (1976) "The Estimated Impact of the Antidiscrimination Program Aimed at Federal Contractors," *ILR review*, 29, 523–543.

Goodman-Bacon, Andrew (2021) "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 225, 254–277.

Heckman, James J. and Kenneth I. Wolpin (1976) "Does the Contract Compliance Program Work? An Analysis of Chicago Data," *ILR Review*, 29, 544–564.

Holzer, Harry J. and David Neumark (2000a) "What Does Affirmative Action Do?" *ILR Review*, 53, 240–271.

Holzer, Harry and David Neumark (1999) "Are Affirmative Action Hires Less Qualified? Evidence from Employer-Employee Data on New Hires," *Journal of Labor Economics*, 17, 534–569.

Holzer, Harry and David Neumark (2000) "Assessing Affirmative Action," *Journal of Economic Literature*, 38, 483–568.

Kline, Patrick, Evan K. Rose, and Christopher R. Walters (2022) "Systemic Discrimination among Large US Employers," *Quarterly Journal of Economics*, 137, 1963–2036.

Kline, Patrick, Evan K. Rose, and Christopher R. Walters (2024) "A Discrimination Report Card," *American Economic Review*, 114, 2472–2525.

Kurtulus, Fidan Ana (2012) "Affirmative Action and the Occupational Advancement of Minorities and Women During 1973–2003," *Industrial Relations: A Journal of Economy and Society*, 51, 213–246.

Kurtulus, Fidan Ana (2016) "The Impact of Affirmative Action on the Employment of Minorities and Women: A Longitudinal Analysis Using Three Decades of EEO-1 Filings," *Journal of Policy Analysis and Management*, 35, 34–66.

Leibbrandt, Andreas and John List (2018) "Do Equal Employment Opportunity Statements Backfire? Evidence from a Natural Field Experiment on Job-Entry Decisions," NBER Working Paper Series.

Leonard, Jonathan S. (1984) "The Impact of Affirmative Action on Employment," *Journal of Labor Economics*, 2, 439–463.

Leonard, Jonathan S. (1996) "Wage Disparities and Affirmative Action in the 1980's," *American Economic Review*, 86, 285–289.

Miller, Conrad (2017) "The Persistent Effect of Temporary Affirmative Action," *American Economic Journal: Applied Economics*, 9, 152–90.

Rodgers, William M. and William E. Spriggs (1996) "The Effect of Federal Contractor Status on Racial Differences in Establishment-Level Employment Shares: 1979-1992," *American Economic Review*, 86, 290–293.

Scikit-learn (2019) "Fuzzy Matching at Scale by Josh Taylor," URL: https://towardsdatascience.com/fuzzy-matching-at-scale-84f2bfd0c536.

Smith, James P. and Finis Welch (1984) "Affirmative Action and Labor Markets," *Journal of Labor Economics*, 2, 269–301.

Sorkin, Isaac (2018) "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 133, 1331–1393.

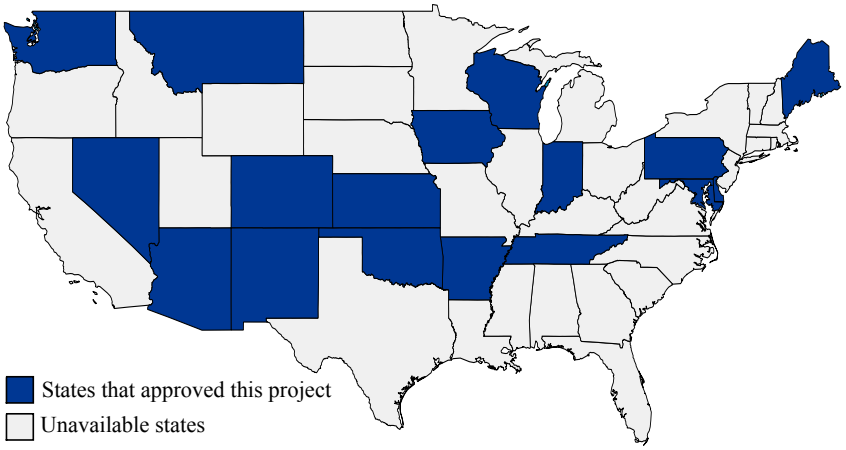
Sun, Liyang and Sarah Abraham (2021) "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 225, 175–199.

Vilhuber, Lars et al. (2018) "LEHD Infrastructure S2014 Files in the FSRDC," *US Census Bureau, Center for Economic Studies Discussion Papers, CES*, 1, 3.

Appendix A. LEHD States

We have access to the Longitudinal Employer-Household Dynamics (LEHD) data spanning from 2001 to 2014[65] for the 19 states displayed in Figure A.1.

Figure A.1: *States analyzed in this paper.*



Notes: The figure shows in blue the states that approved this project and therefore gave us access to their data: Arizona, Arkansas, Colorado, District of Columbia, Delaware, Hawaii, Indiana, Iowa, Kansas, Maine, Maryland, Montana, Nevada, New Mexico, Oklahoma, Pennsylvania, Tennessee, Washington, and Wisconsin.

The LEHD comprises three datasets: the Employer Characteristics File (ECF), the Employment History File (EHF), and the Individual Characteristics File (ICF). We linked them together following Vilhuber et al. (2018) and Sorkin (2018) through internal identifiers defined by the US Census Bureau.

The ECF consolidates most employer and establishment-level information, including size, age, location, and industry classification. The EHF documents comprehensive in-state employment histories for individuals, represented through employer-employee linkages, and includes data on quarterly labor earnings.[66]

The ICF complements these datasets with socio-demographic details of the workforce, such

[65] While more recent years of the LEHD were available at the time of writing this paper, we did not apply for access to the LBD or SSEL after 2012. These datasets are key in linking together the LEHD with the FPDS.

[66] Earnings data, as captured here, include "gross wages and salaries, bonuses, stock options, tips and other gratuities, and the value of meals and lodging" and omit "employer contributions to Old-age, Survivors, and Disability Insurance (OASDI); health insurance; unemployment insurance; workers' compensation; and private pension and welfare funds."

as age, gender, race,[67] ethnicity,[68] birthplace, and citizenship status.

Although coverage by the LEHD varies by state, it covers, on average, 96% of all private-sector jobs in the U.S. More information about the LEHD data is available in Abowd et al. (2009) and Vilhuber et al. (2018).

To link the LEHD files with the rest of the data sources we use, we integrate them with the Standard Statistical Establishment Listing (SSEL) and the Longitudinal Business Database (LBD), adding establishment names and addresses.[69] See Figure B.1 for an illutration of the linkages we made. This linkage is available from 2001 to 2012 which makes this the period in which we study firms becoming subject to affirmative action.

Appendix B. Data Linkages

Given the size of our core data, a key challenge was merging the some of our data sources we use without a reliable common identifier, specifically, the FPDS and Audit records to the LEHD. To overcome this, we modify a fuzzy matching method based on the term frequency-inverse document frequency (tf-idf) algorithm (Scikit-learn, 2019) to account for two dimensions: establishment name and address and we merge the FPDS and Audit records to the SSEL, which has common identifiers in the LEHD.

The tf-idf score quantifies a word's importance relative to a document within a corpus (a collection of documents). This score increases with the word's frequency in a document but is offset by its frequency across all documents, mitigating the impact of commonly occurring words.[70]

Our method leverages cosine similarity between tf-idf vectors to measure the similarity between establishment names and addresses. This approach allows for efficient comparison of text strings, crucial for our large-scale data merging task.

We split the U.S. into 3-digit zipcode regions and within each of these regions we proceed as follows.

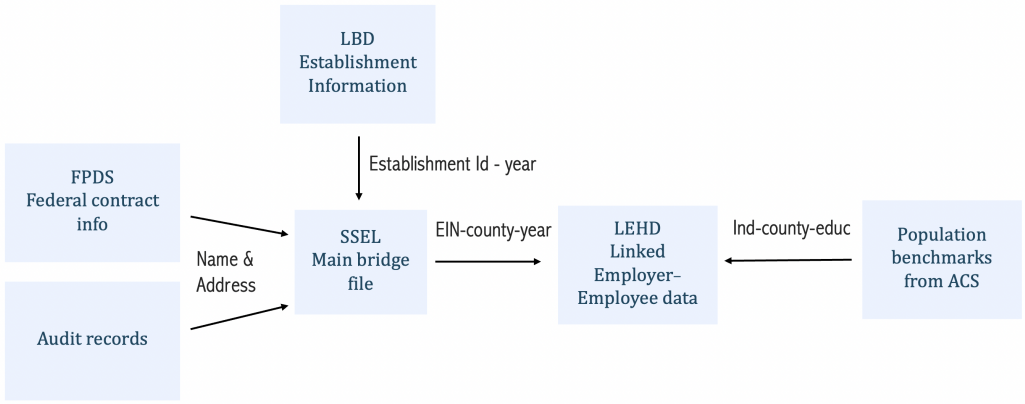
[67] Race is categorized into: 1) White Alone, 2) Black or African American Alone, 3) American Indian or Alaska Native Alone, 4) Asian Alone, 5) Native Hawaiian or Other Pacific Islander Alone, and 7) Two or More Races.

[68] Hispanic or Latino identity is flagged separately in the ICF.

[69] While both the FPDS and OFCCP's audit data have Dun & Bradstreet (DUNS) numbers (a unique nine-digit identifier for businesses), these codes are often incomplete or missing. Therefore, without names and addresses, which are rarely missing, we would not be able to link our data.

[70] Though we considered alternative string similarity measures like Levenshtein or Jaro-Winkler distances, these alternatives are not suited for dealing with large data as the time required to implement them grows quadratically with the number of observations.

Figure B.1: *Data linkages to construct our novel data.*



Notes: The diagram displays the data linkages we make to construct our novel data. The FPDS and Audit records are linked to the SSEL via the two-dimensional fuzzy matching method described in this section. The SSEL has common identifiers in the LEHD.

1. We start by considering the names of establishments and split these into 3-character n-grams.[71] We create tf-idf vectors for names' n-grams in the FPDS (training corpus) and LEHD (query corpus). The tf-idf scores are normalized so that they have norm 1 and they belong to $\mathbb{R}^K$ where K denotes the total number of distinct n-grams in the training corpus.[72] Then, we compute the cosine similarity between the vectors from both datasets, which is the angle in between the tf-idf scores' vectors,[73] and identify potential matches based on a similarity threshold.

2. Then, for all potential matches based on name (those with a cosine similarity score above the threshold), we repeat the process with establishment addresses to confirm that all matches meet both name and address similarity criteria.

Our matching algorithm efficiently handles large datasets unlike traditional string distance measures, which can be resource-intensive, making it scalable for extensive data integration tasks. It easily accounts for minor spelling variations and typos, enhancing the accuracy of matches, and can be flexibly tailored applications for different datasets.

[71] That is, we split the name of each establishment into 3-character pieces. As an example, the n-grams of the word 'Department', would be ' De', 'Dep', 'epa', 'par', 'art', 'rtm', 'tme', 'men', 'ent', 'nt '.

[72] If any of the names in the query corpus have no n-grams in common with the names in the training corpus, the name is mapped into a vector of zeros, $0^K \in \mathbb{R}^K$. Otherwise, these K dimensional vectors have norm 1.

[73] The cosine similarity score between vectors $x, y \in \mathbb{R}^K$ is given by $x \cdot y / \|x\| \|y\| = \cos(x, y)$. That is, it measures the angle between x and y, where the angle (with range $(0, \pi)$) is mapped into the $(0, 1)$ interval.

B.1 EEO-Tabulations

The EEO Tabulations[74] are constructed based on 5-year ACS data by the Census Bureau in consultation with the EEOC and other members of the EEO Tabulation Consortium. These tabulations are available at the location-industry-occupation level, however, the LEHD contains no information on the workers' occupations.

While these tabulations are publicly available, we cannot directly link them to the LEHD, as these data lack occupations. Thus, we reconstruct workforce tabulations by county using the same 5-year ACS data employed by the Census, to approximate the industry- and occupation-level proportions in the EEO-Tabulations using racial, ethnic, and gender shares by industry across broad education levels: college and non-college. These reconstructed tabulations allow us to define an approximate target share of Black workers, which we link to our data. While not perfect proxies, our tabulations provide a useful approximation to EEO-Tabulations.

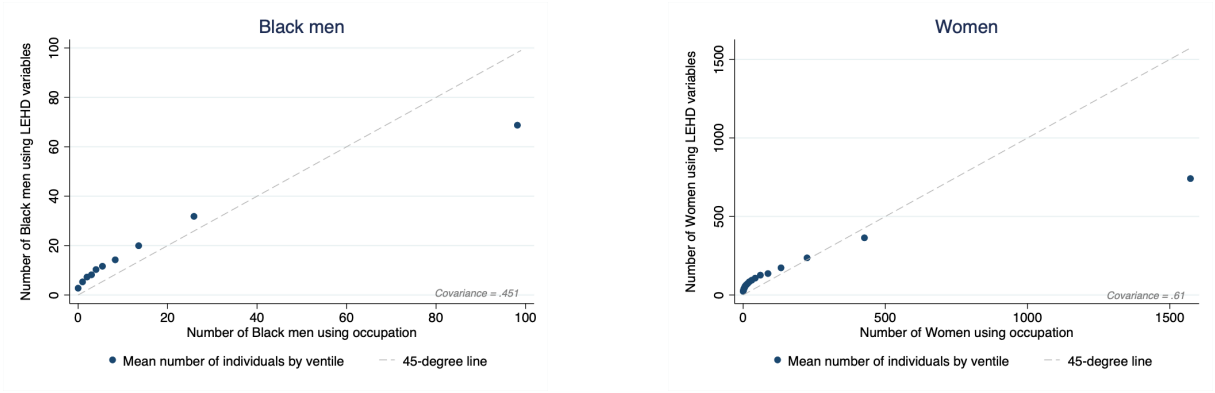
B.1.1 Tabulations' comparison using education to proxy for occupation

We use the ACS 5-year extract from 2006-2010 and aggregate the number of workers by location-industry across broad education levels: college and non-college.

The following binned scatterplots illustrate the correlation between the two measures. We refer to the numbers computed using by location-industry-occupation as the *occupation counts* and those obtained by location-industry-education-age as the *education-age counts*. The plots are constructed by grouping the occupation counts into ventiles. Then we compute the mean of each of our counts within each of these bins. The figures show the scatterplot of these data points. For reference, we also plot the identity line (which would underlie the dot in that bin if the means of the counts coincided).

[74] Available at https://www2.census.gov/EEO_2006_2010/EEO_2006_2010_Tables_Nation(010)_CSV/.

Figure B.2: *Correlation between the EEO-tabulations and our estimates*



(a) *Black men* ***(b)*** *Women*

Notes: The plots are constructed by grouping the occupation counts into equally sized bins, specifically into ventiles. Then we compute the mean of each of our counts within each of these bins and the figures show the scatterplot of these data points. For reference, we also plot the identity line (which would underlie the dot in that bin if the means of the counts coincided).

Appendix C. The Discontinuity of the Policy

We leverage the \$50,000 value threshold in a regression discontinuity design (RDD) to further validate our findings. This approach complements our main specification by exploiting a different source of variation in the policy's application. We impose additional sample restrictions, focusing on contractor firms with 50 or more employees and establishments entering our data with a maximum contract value in the \$25,000–\$75,000 range. To maintain consistency with our staggered event study designs, we exclude establishments that enter our data as contractors and require that contractors be first observed in the specified bandwidth from 2004 onward.

Our RDD estimation equation is as follows:

$$y_{jt}^{k} = \alpha^{k} + \beta^{k} \mathbb{1}_{c_j \geqslant \$50{,}000} + \delta^{k} c_j + \zeta^{k} \mathbb{1}_{c_j \geqslant \$50{,}000} \times c_j + \gamma^{k} X_j + \lambda_t + \varepsilon_{jt} \tag{3}$$

where y_{jt}^{k} is the outcome of interest k periods after establishment j crosses the contract value threshold, $\mathbb{1}_{c_j \geqslant \$50{,}000}$ indicates treatment status, X_j is a vector of pre-regulation controls, and λ_t represents year fixed effects. For multi-contract establishments, c_j is the highest contract value in a given year.

We focus on two key outcomes: the percentage of Black workers (stock measure) and the percentage of Black new hires (flow measure), evaluated at treatment year ($k = 0$) and one year later. Table C.1 presents the qualitative results from estimating β^{k} in equation (3), with standard

errors clustered at the firm level.

Table C.1: *Qualitative estimates from* (3)

	Proportion Black workers			Proportion Black new hires		
$t+k$, $k=$	-1	0	2	-1	0	2
β^k	+	−	−	+	−	+
ζ^k	+	+	−	+	−	−
δ^k	−	+	+	−	+	−
γ^k						
firm age	−	−	+	-	+	−
Firm employment	$+^{*}$	+	+	$+^{**}$	$+^{*}$	+
Constant	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$	$+^{***}$
Number of observations	15,700	17,600	15,500	13,000	15,700	14,000

*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Notes: This table presents the qualitative results for the estimates of β^k in model (3), for $k \in \{-1,0,2\}$. $-$ indicates a negative point estimate, $+$ a positive one. N is the number of establishments, rounded to the nearest 100.

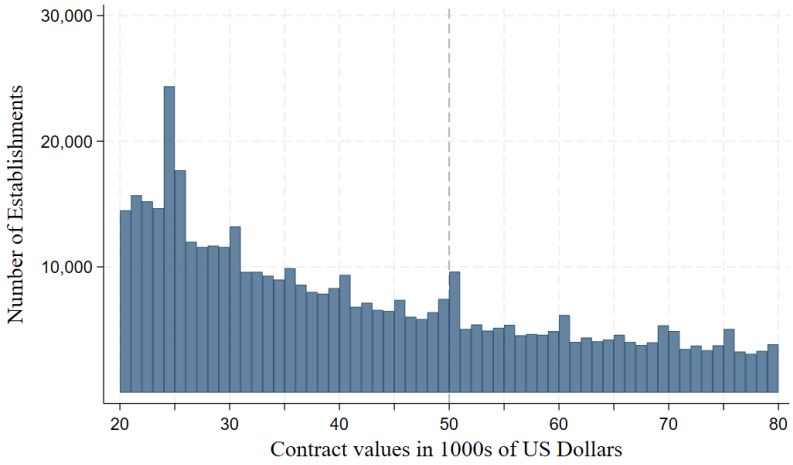
Crucially, we find no statistically significant pre-treatment differences ($\hat{\beta}^{-1}$), supporting the validity of our RDD. Moreover, consistent with our main findings in Section 5, we detect no significant impact of the affirmative action plan on any outcomes: estimates $\{\hat{\beta}^k\}_{k\in\{0,1\}}$ are statistically insignificant and fluctuate around zero.

These RDD results corroborate our main findings, providing robust evidence that affirmative action policies, as implemented during our study period, did not significantly affect the share of Black workers, the share of Black new hires, or the Black-white earnings gap. This consistency across methodologies strengthens our conclusion about the limited effectiveness of these policies in achieving their intended diversity goals, underscoring the need for a critical reassessment of affirmative action implementation in federal contracting.

Appendix D. Evidence Supporting our Identifying Assumption

Selection into federal contracting could threaten our identifying assumption. If firms with lower shares of minority workers perceive the compliance cost of affirmative action as higher than other firms and deliberately seek contracts under $50,000 or not bid for contracts at all. This behavior would violate our identification assumption. However, if this was the case, we would expect to see that the value of federal contracts cluster right below $50,000. Figure D.1 shows that this does not seem to be happening: If anything, there is a cluster of contracts just *above* the threshold.

Figure D.1: *No bunching below the $50,000 threshold.*



Notes: This figure plots a histogram of contract values across establishment-years in our data. The sample includes 228,199 distinct establishments, some of which receive multiple contracts from 2001 to 2017, resulting in 447,521 establishment-year observations.

Table D.1: *Contracts' characteristics above and below the $50,000 threshold.*

	All contracts	**Between $25-50K**	**Between $50-75K**
Contract characteristics			
Duration (years)	1.59	1.62	1.64
Competed	0.68	0.69	0.68
Fixed price	0.89	0.91	0.89
Offers received	3.85	3.93	4.08
Awarding Agency			
Dept. of Defense	0.57	0.58	0.59
Gral. Services Administration	0.11	0.13	0.11
Veteran Affairs	0.09	0.08	0.07
Other	0.23	0.22	0.24
Number of contracts	24,601,034	4,044,294	1,768,704

Notes: The table shows in the top panel the average characteristics of the federal contracts from 2001 to 2017. In particular the duration they have, the share that was awarded through an open competition, the share with a fixed price charged to the government, and the number of actual offers/bids received in response to the solicitation. The bottom panel shows the awarding agency, where *other* includes roughly 70 other agencies, among which the majority of the contracts in this group belong to the following agencies: Department of Justice, Department of Health and Human Services, Department of the Interior, Department of Agriculture, Department of State, Department of Homeland Security, Department of Transportation, Department of the Treasury, Department of Commerce, National Aeronautics and Space Administration, Environmental Protection Agency, Social Security Administration, Office of Personnel Management, and Department of Energy.
This excludes contracts worth $10,000 or less, which are exempt from EO 11246, and those worth $10 million or more, which require a pre-award compliance evaluation that inspects the racial composition of the contractor.

Appendix E. EEO-1 Reports

We got access to the EEO-1 records collected by the EEOC for the years 1990–2014 at an earlier stage of the project. All the results produced with EEO-1 records were generated when we had access to the data.

The EEO-1 records are comprised by establishment-level data in which each firm reports their overall racial,[75] and gender composition of their personnel, disaggregated by 9 major occupation groups.

The EEOC requests firms to fill EEO-1 records, if they larger than 100 workers unless they are federal contractors, in which case the size requirement goes down to 50 workers.[76] Firms are required to file a separate report for its headquarters, and for each establishment of at least 50 employees. We observe a firm and establishment identifier, each establishment location and industry, its size, and its contractor status. Importantly, the contractor and regulation status of each establishment is self-reported in these records.[77]

Table E.1: *EEO-1 Reporting Rates by Industry*

Industry	**EEO1 Coverage Rate (%)**	**Industry Size (%)**
Agricultural Services	21.2	0.5
Mining	71.0	0.5
Construction	10.1	5.8
Manufacturing	53.8	20.0
Transportation, Communications, Utilities	50.3	6.1
Wholesale Trade	18.2	6.9
Retail Trade	25.3	21.0
Finance, Insurance, and Real Estate	40.7	7.8
Services	28.9	31.3
Overall	33.7	100.0

Notes: The table shows coverage rates for the EEO-1 reports from 1990 to 2014. Coverage rates are calculated by dividing EEO-1 reported employment by County Business Patterns employment totals for the 204 MSAs used in the analysis. Industry size is the fraction of total County Business Patterns reported employment in that industry.

Due to size requirements, EEO-1 data is not representative of all U.S. establishments. Table E.1 tabulates coverage rates for this data by comparing it to County Business Patterns data.

[75] There is no distinction between race and ethnicity in the data. In particular, Hispanic workers are classified as a distinct, non-overlapping group.

[76] These size requirements were adjusted in 1982. Before then, the requirement was 50 workers for all firms, and 25 workers for federal contractors.

[77] In the EEO-1 record filed to the EEOC, the establishment reports whether it is a contractor, and whether it is subject to EO 11246 regulation.

Industries that tend to have larger establishments are overrepresented (e.g. mining, manufacturing), and industries that tend to have small establishments are underrepresented (e.g. services). Overall, EEO-1 data accounts for about 34% of total employment.

E.1 Benchmarking our findings to Miller (2017)

We compare our findings with those of Miller (2017) in Section 5.3 using EEO-1 records. Here, we delve into the reporting patterns observed in these records, which are detailed in Section E.1.1. We then present additional evidence suggesting that the discrepancies between our null results and those of Miller (2017) are primarily attributable to the different time periods we examine in our respective studies.

E.1.1 Discrepancies in reported affirmative action statuses

While the effects we find with our data and with the EEO-1 records are consistent and suggest that the positive effects we found have diminished over time, the results from the EEO-1 records trend positive (though insignificant) compared with our main results.

We argue that this is largely due to the differences in the differences between how the regulation status in the EEO-1 records is measured, as compared to the FPDS, the source for our novel data.

The EEOC requires annual EEO-1 reports from companies with 100 or more employees and federal contractors employing 50 or more workers. These reports consist of self-reported information every year about 'affirmative action status' and workforce diversity. Establishments of the firm with at least 50 employees are required to file a separate report. The specific question that EEO-filers answer regarding 'affirmative action status' is whether the company or any of its establishments have 50 or more employees and have a federal contract (or subcontract) of $50,000 or more.[78] Table E.2 reveals that, on average, 6% of firms report different answers for this question across their establishments. Among these firms with mixed reporting, 43% of their establishments report that neither the company nor other establishments are executing a federal contract of over $50,000, while the remaining establishments, report the opposite. While these reporting patterns are allowed and are not indicative of a violation, they suggest that some establishments may be unaware of their regulatory obligations. Furthermore, since the reporting requirement applies only to establishments meeting specific size criteria that differs by federal

[78] See Question C3 of the sample EEO-1 form at http://www.eeo1.com/EEO1Form.pdf.

Table E.2: *Share of firms and establishments with discrepancies in their reported affirmative action status in EEO-1 forms.*

Year	Num. Firms	% Firms with misreports	% Establishments misreporting	
			Across all firms	Among misreporting firms
1990	38,698	4.79	14.10	50.30
1991	38,372	3.94	10.89	47.15
1992	38,220	4.20	11.89	47.05
1993	38,513	4.28	10.51	50.20
1994	39,138	4.17	10.37	48.05
1995	39,905	4.52	11.17	50.50
1996	39,893	4.63	11.89	53.67
1997	39,680	3.83	9.50	47.08
1998	39,666	3.51	9.00	43.85
1999	38,675	6.63	12.89	50.69
2000	39,482	5.06	9.97	50.90
2001	39,350	4.68	9.67	48.75
2002	39,839	4.20	7.66	41.91
2003	40,059	4.91	7.21	41.64
2004	42,628	4.81	6.96	44.28
2005	46,601	20.46	10.02	31.16
2006	49,501	8.20	10.49	42.70
2007	67,848	17.77	9.17	36.93
2008	68,369	6.33	9.59	41.37
2009	67,143	7.60	10.21	40.55
2010	67,449	7.97	10.66	42.55
2011	67,971	7.38	10.86	45.27
2012	68,316	1.96	4.44	37.05
2013	69,252	1.82	5.12	42.80
2014	67,152	0.00	0.00	·
Average	49269	5.91	9.77	43.06

Notes: The table shows the yearly number of firms filing EEO-1 forms to the EEOC. The percentage of firms with inconsistencies in their reported affirmative action status across establishments ranges from 2% to 20%. Overall, about 10% of establishments misreport their affirmative action status when other units of the same company report that they are under affirmative action. This percentage ranges from 30% to 60% among firms with discrepancies in their affirmative action status.

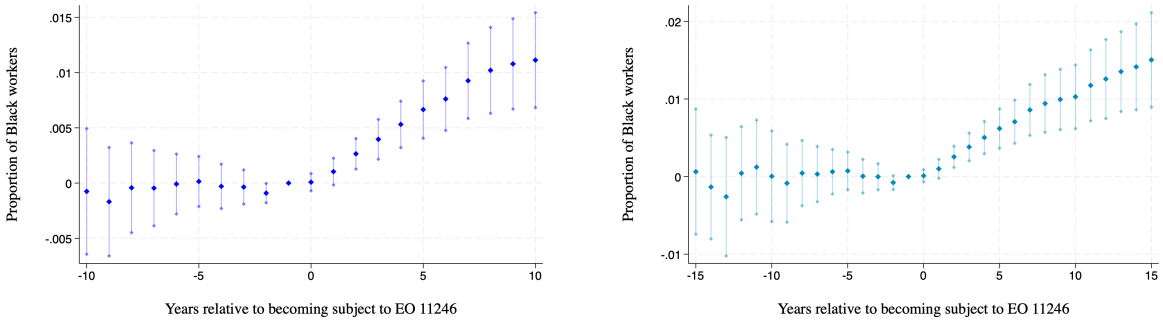
contractor status, some eligible establishments could inadvertently fail to report, mirroring the reporting discrepancies observed in submitted forms and, therefore, contributing to selection.

E.1.2 Replicating Miller (2017) with more leads and lags

The diminishing effects we find with the EEO-1 reports could be due to diminishing effects of affirmative action across the board or to newly regulated firms responding differently from those that were regulated in the 90s. Replicating Miller's (2017) main specification using 10 leads and lags, may suggest that affirmative action's effects have diminished over time across the board; also for firms that were regulated in the 90s. However, further extending the period of analysis to include 15 leads and lags adds nuance to this picture.

Figure E.1: *EEO-1 Reports: Share of Black Workers Upon Becoming Subject to EO 11246*

(a) *Event Study with 10 leads and lags* ***(b)*** *Event Study with 15 leads and lags*



Notes: These figures show event study estimates using EEO-1 Reports from 1990 to 2014. ***Panel (a)*** displays event study estimates of the leads and lags for an establishment first becoming subject to affirmative action, where we include 10 leads and lags. ***Panel (b)*** shows an analogous graph where we include 15 leads and lags. Bars represent 95% confidence intervals. We normalize $\theta_{-1} = 0$ and cluster standard errors at firm level.

E.1.3 Replicating Miller (2017) while considering only the states in our data

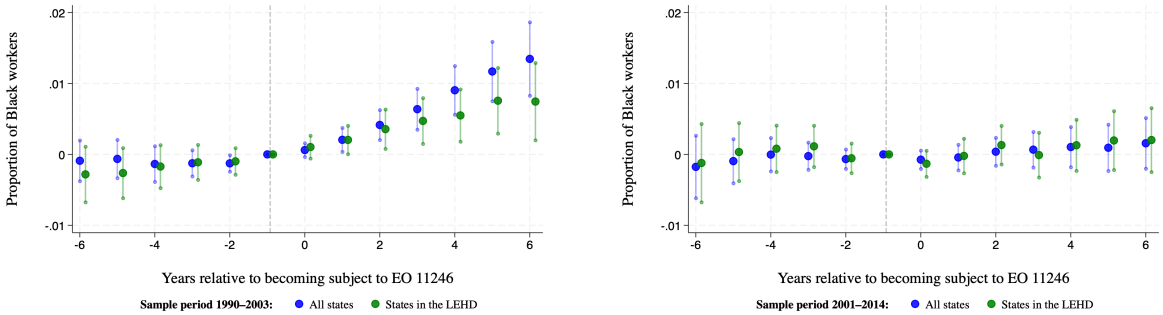
To ensure comparability of the results we present in Section 5.3, with our novel dataset, we restrict our analysis to the subset of states represented in our primary data (see Figure A.1).

We begin by replicating the main results from Miller (2017) using the same approach as in Figure 4 (a), employing 14-year time windows. In our EEO-1 reports sample, these windows range from the earliest period of 1990–2003 to our primary analysis period of 2001-2014. Figure E.2 presents estimates for both the full sample of establishments and the subset corresponding to states in our novel dataset for the first and last of these periods (we present the set of estimates for all time periods in Figure 4 (a) and, for our set of states, in Figure E.4).

While the positive effects observed in the 1990–2003 period are less pronounced in our subset of states, they remain positive and statistically significant. In contrast, for the 2001–2014 period, both sets of estimates are statistically indistinguishable from zero.

Figure E.2: *Main event study from Miller (2017) in the EEO-1 Reports and in the subsample of states available in the LEHD*

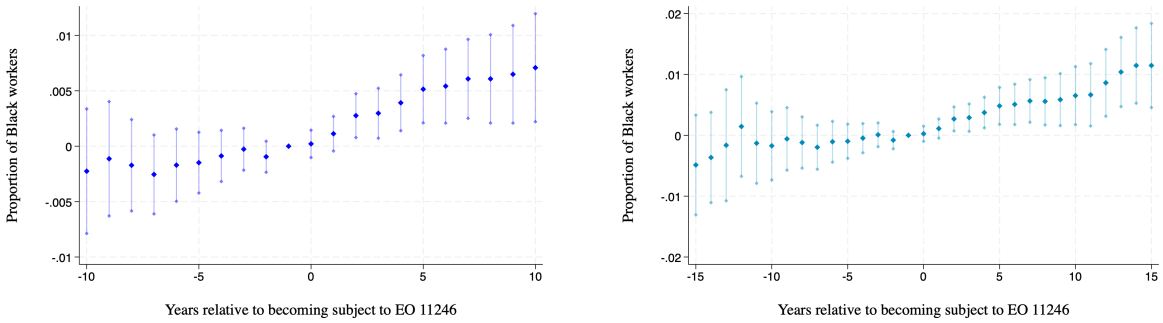
(a) *Event Study in 1990–2003* ***(b)*** *Event Study in 2001–2014*



Notes: These figures show event study estimates using EEO-1 Reports from 1990 to 2014. ***Panel (a)*** restricts the time period of analysis to 1990–2001 and displays event study estimates of the leads and lags for an establishment first becoming subject to affirmative action, in blue, and compares these to the analogous estimates requiring establishments to be in the states that appear in the LEHD. ***Panel (b)*** shows estimates the corresponding estimates for the period 2001–2014, which coincides with the period of our novel data. Bars represent 95% confidence intervals. We normalize $\theta_{-1} = 0$ and cluster standard errors at firm level.

Figure E.3: *Adding leads and lags to the main event study from Miller (2017) in the states in our data*

(a) *Event Study in 1990–2014* ***(b)*** *Event Study in 2001–2014*



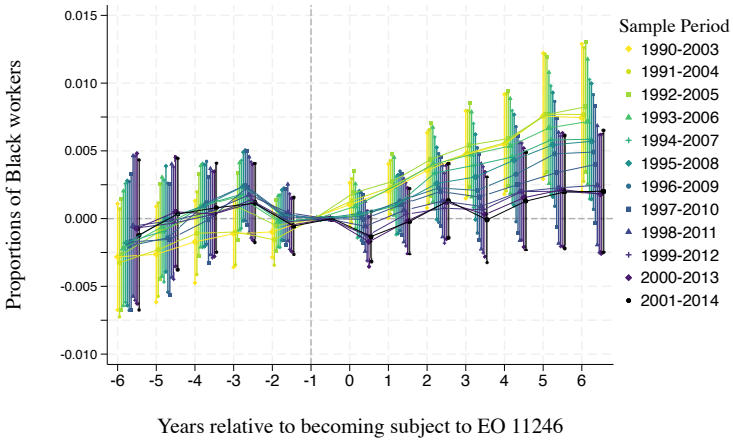
Notes: These figures show event study estimates using EEO-1 Reports from 1990 to 2014 for the subset of states that appear in our novel data. ***Panel (a)*** displays event study estimates of the leads and lags for an establishment first becoming subject to affirmative action, where we include 10 leads and lags. ***Panel (b)*** shows an analogous graph where we include 15 leads and lags. Bars represent 95% confidence intervals. We normalize $\theta_{-1} = 0$ and cluster standard errors at firm level.

Figure E.3 presents a replication of the event study design used by Miller (2017), but with an expanded set of leads and lags. This analysis mirrors Figure E.1, focusing specifically on our subset of states. The results reveal a pattern similar to that observed in Figure E.1: while the positive effects in our set of states are less pronounced, they remain statistically significant and follow a comparable trend. This consistency across different state subsets reinforces the

robustness of our findings and provides additional context for interpreting the results within our novel data.

Lastly, we replicate Figure 4 but for our subset of states and continue to see that while the positive effects in the earlier periods are less prominent, they continue to diminish over time.

Figure E.4: *Event Study Estimates Over Time in LEHD states*

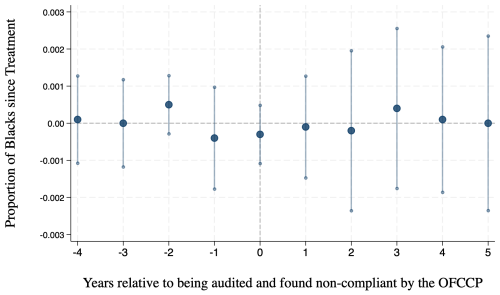


Notes: These figures show event study estimates using EEO-1 Reports from 1990 to 2014 for the set of states in our novel data, where each line displays event study estimates of the leads and lags for an establishment first becoming subject to affirmative action for one of 12 time periods. The earliest period starts in 1990 and ends in 2003, and the latest corresponds to our study period with novel data from 2001 to 2014. Each time window spans 14 years, but the *first* time an establishment is defined to be under affirmative action, it is restricted to not having been regulated since 1990, the first available year.

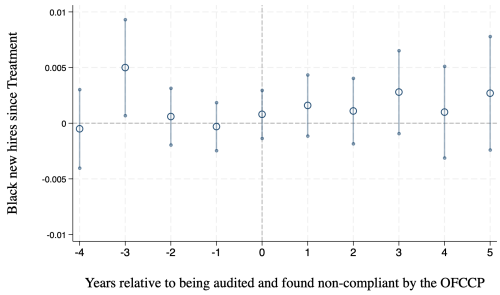
Appendix F. The Effect of Findings of Non-Compliance

Figure F.1: *Worker-level outcomes upon an audit for the set of establishments that do not get a notice of compliance of affirmative action*

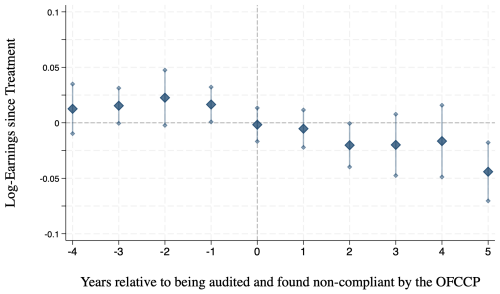
(a) *Black workers' log-earnings*



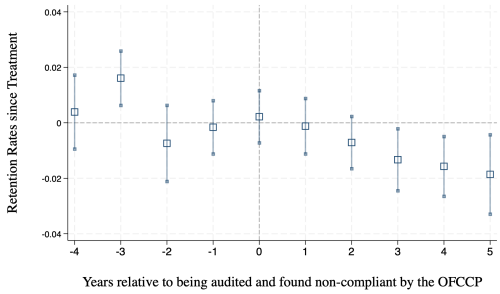
(b) *Black workers' retention rates*



(c) *Black workers' log-earnings*



(d) *Black workers' retention rates*



Notes: The figure plots the weighted average of our estimated ATTs. These estimates capture the average dynamic impact of EO 11246 measured e years after being audited, two worker-level outcomes. **Panel (a)** shows the estimated dynamic effects of the treatment on the share of Black workers in audited establishments that failed the audit. The sample size rounded to the nearest 1,000 used for estimation is $165,000$. **Panel (b)** shows the estimated effects of the treatment on the share of Black new hires in audited establishments that failed the audit. The sample size rounded to the nearest 1,000 used for estimation is $154,000$. **Panel (c)** shows the estimated dynamic effects of the treatment on the log-earnings of Black workers in audited establishments. The sample size rounded to the nearest 1,000 used for estimation is $881,000$. **Panel (d)** shows the estimated effects of the treatment on the retention rates of Black workers in audited establishments. The sample size rounded to the nearest 1,000 used for estimation is $873,000$. The bars indicate the 95% confidence intervals. We cluster standard errors at the firm level.

Appendix G. OFCCP Data

G.1 OFCCP's Historical Budget

Next, we present the Congressional appropriation for the OFCCP, meaning the budget allocated by Congress to the OFCCP. As Table G.1 shows, The budget used to increase by about 5% year

by year before the Obama administration, when it increased by 27%. Since then, the budget increased on average by only 2% until 2022, when it increased by 11%.

Table G.1: *Historical OFCCP's budget and full time equivalent (FTE) workers*

Presidential Administration	**Year**	**FTE**	**Allocated Budget** (1000's Dollars)	**2010 adjusted** (1000's Dollars)	**% Change** from previous year
Clinton	1993				
	1994	785	56,443	38,353	
	1995	775	58,928	41,177	7.36%
	1996	727	56,851	40,898	-0.68%
	1997	712	59,058	43,461	6.27%
	1998	743	62,271	46,539	7.08%
	1999	727	65,461	50,004	7.44%
	2000	811	73,250	57,834	15.66%
Bush	2001	776	76,000	61,713	6.71%
	2002	718	77,701	64,092	3.85%
	2003	742	78,033	65,833	2.72%
	2004	749	79,441	68,805	4.52%
	2005	691	80,059	71,690	4.19%
	2006	670	81,285	75,136	4.81%
	2007	625	82,442	78,360	4.29%
	2008	585	81,001	79,961	2.04%
Obama	2009	596	84,172	82,783	3.53%
	2010	788	105,386	105,386	27.30%
	2011	775	105,386	108,672	3.12%
	2012	755	105,187	110,733	1.90%
	2013	729	99,685	106,495	-3.83%
	2014	683	104,976	113,929	6.98%
	2015	621	106,476	115,703	1.56%
	2016	615	105,476	116,067	0.31%
Trump	2017	556	104,476	117,410	1.16%
	2018	514	103,476	119,133	1.47%
	2019	500	103,476	121,315	1.83%
	2020	472	105,976	125,752	3.66%
Biden	2021	496	105,976	131,680	4.71%
	2022	451	108,476	145,580	10.56%
	2023	495	110,976	155,041	6.50%
	2024	492			

Notes: FTE denotes full time equivalent workers. The allocated budget refers to the funds that Congress allocates to a government agency — this is the *Congressional Appropriation* and not the *Presidential Request*. The FTE and budget data from 1994–2002 were obtained from the President's 2006 Request for Funding Federal Civil Rights Enfocement, which uses figures from the Department of Labor. The analogous figures from 2003 onward were obtained from the Congressional Budget Justification for the fiscal years of 2011, 2013, 2014, 2021, and 2025.

G.2 Violations found by the OFCCP

GAO (2016) provides a comprehensive assessment of the enforcement measures that the OFCCP takes. While the data used by the Government Accountability Office generally follows the same trend that our audit records, we have a few more audits in our data and present the discrepancies in Table G.2.

Table G.2: *Types of Violations Cited by the OFCCP as a Result of Compliance Evaluations in Fiscal Years 2010–2015*

Fiscal Year	Total cases	No violations	Violations (%) All	Violations (%) Discrimination	Violations (%) Technical
	(1)	(2)	(3)	(4)	(5)
2004	6570	5964 (90.78)	606 (9.22)	79 (1.20)	573 (8.72)
2005	2727	2292 (84.05)	435 (15.95)	66 (2.42)	409 (15.00)
2006	3988	3426 (85.91)	562 (14.09)	78 (1.96)	522 (13.09)
2007	4860	4251 (87.47)	609 (12.53)	74 (1.52)	586 (12.06)
2008	4255	3566 (83.81)	689 (16.19)	95 (2.23)	652 (15.32)
2009	3848	3108 (80.77)	740 (19.23)	96 (2.49)	688 (17.88)
2010	4916	3886 (79.05)	1030 (20.95)	101 (2.05)	971 (19.75)
2011	3925	2733 (69.63)	1192 (30.37)	158 (4.03)	1096 (27.92)
2012	3956	2534 (64.05)	1422 (35.95)	186 (4.70)	1318 (33.32)
2013	4071	2861 (70.28)	1210 (29.72)	144 (3.54)	1119 (27.49)
2014	3833	3196 (83.38)	637 (16.62)	87 (2.27)	577 (15.05)
2015	2602	2100 (80.71)	502 (19.29)	49 (1.88)	469 (18.02)
2016	1696	1347 (79.42)	349 (20.58)	50 (2.95)	327 (19.28)
2017	1142	885 (77.50)	257 (22.50)	46 (4.03)	233 (20.40)
2018	812	647 (79.68)	165 (20.32)	48 (5.91)	146 (17.98)
2019	1343	1173 (87.34)	170 (12.66)	93 (6.92)	122 (9.08)
2020	1320	1313 (99.47)	7 (0.53)	6 (0.45)	6 (0.45)
Average %		(81.37)	(18.63)	(2.97)	(17.11)

Notes: The table presents the types of violations identified by the OFCCP during compliance evaluations for fiscal years 2004–2020 — the period that our audit data cover. We categorize all violations into two types: technical ones and discrimination findings. Technical violations include failures in record keeping, omission of the EEO-tag line on job ads, or issues with the AAP. Discrimination findings encompass issues related to hiring, firing, promotion, and salary decisions; violations in selection, testing, or medical screening procedures; failures to accommodate workers; and instances of systemic discrimination. All other violations in the data, such as "past performance," are neither classified as discrimination findings nor technical violations but are still counted as violations. Threfore, these are counted in column (3) but not in (4) or (5). Moreover, a single establishment can be found to have both technical and non-technical violations; therefore, the sum of columns (4) and (5) may not equal column (3).

Appendix H. Executive Order 11246

The following is the text from Sections 202 and 203 of the original Executive Order 11246, which outlines the nondiscrimination and affirmative action obligations of government contractors.[79]

> Sec. 202. Except in contracts exempted in accordance with Section 204 of this Order, all Government contracting agencies shall include in every Government contract hereafter entered into the following provisions: During the performance of this contract, the contractor agrees as follows:
>
> 1. The contractor will not discriminate against any employee or applicant for employment because of race, color, religion, sex, or national origin. The contractor will take affirmative action to ensure that applicants are employed, and that employees are treated during employment, without regard to their race, color, religion, sex, or national origin. Such action shall include, but not be limited to the following: employment, upgrading, demotion, or transfer; recruitment or recruitment advertising; layoff or termination; rates of pay or other forms of compensation; and selection for training, including apprenticeship.
>
> 2. The contractor will, in all solicitations or advertisements for employees placed by or on behalf of the contractor, state that all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, or national origin.
>
> 3. The contractor will send to each labor union or representative of workers with which he has a collective bargaining agreement or other contract or understanding, a notice to be provided by the agency contracting officer, advising the labor union or workers' representative of the contractor's commitments under Section 202 of Executive Order No. 11246 of September 24, 1965, and shall post copies of the notice in conspicuous places available to employees and applicants for employment.
>
> *(Sec. 202 continues below...)*

[79] See https://archives.federalregister.gov/issue_slice/1965/9/28/12315-12325.pdf.

4. The contractor will comply with all provisions of Executive Order No. 11246 of September 24, 1965, and of the rules, regulations, and relevant orders of the Secretary of Labor.

5. The contractor will furnish all information and reports required by Executive Order No. 11246 of September 24, 1965, and by the rules, regulations, and orders of the Secretary of Labor, or pursuant thereto, and will permit access to his books, records, and accounts by the contracting agency and the Secretary of Labor for purposes of investigation to ascertain compliance with such rules, regulations, and orders.

6. In the event of the contractor's noncompliance with the nondiscrimination clauses of this contract or with any of such rules, regulations, or orders, this contract may be canceled, terminated, or suspended in whole or in part and the contractor may be declared ineligible for further Government contracts in accordance with procedures authorized in Executive Order No. 11246 of September 24, 1965, and such other sanctions may be imposed and remedies invoked as provided in Executive Order No. 11246 of September 24, 1965, or by rule, regulation, or order of the Secretary of Labor, or as otherwise provided by law.

7. The contractor will include the provisions of paragraphs (1) through (7) in every subcontract or purchase order unless exempted by rules, regulations, or orders of the Secretary of Labor issued pursuant to Section 204 of Executive Order No. 11246 of September 24, 1965, so that such provisions will be binding upon each subcontractor or vendor. The contractor will take such action with respect to any subcontract or purchase order as the contracting agency may direct as a means of enforcing such provisions, including sanctions for noncompliance: Provided, however, that in the event the contractor becomes involved in, or is threatened with, litigation with a subcontractor or vendor as a result of such direction by the contracting agency, the contractor may request the United States to enter into such litigation to protect the interests of the United States.

(Sec. 203 is on the next page)

“Sec. 203. *(a)* Each contractor having a contract containing the provisions prescribed in Section 202 shall file, and shall cause each of his subcontractors to file, compliance reports with the contracting agency or the Secretary of Labor as may be directed. Compliance reports shall be filed within such times and shall contain such information as to the practices, policies, programs, and employment policies and employment statistics of the contractor and each subcontractor, and shall be in such form, as the Secretary of Labor may prescribe.

(b) The Secretary of Labor may investigate the employment practices of any contractor or subcontractor to determine whether or not the contractual provisions specified in Section 202 have been violated. Such investigation shall be conducted in accordance with the procedures established by the Secretary of Labor.

(c) The Secretary of Labor may receive and investigate complaints by employees or applicants for employment with a Government contractor or subcontractor which allege discrimination contrary to the contractual provisions specified in Section 202.”